Profiling the genomic landscape and evolutionary history of polyploid giant cancer cells in undifferentiated pleomorphic sarcomas

Amy L. Bowes, Sara Waise, Tom Lesluyes, Thomas Butters, Christie English, Haixi Yan, Annelien Verfaillie, Christopher Davies, Jianan Chen, Emma Nye, Richard Stone, Jiten Manji, Adrienne M. Flanagan, Jonas Demeulemeester, Maxime Tarabichi, Nischalan Pillay, Peter Van Loo

PII: S0304-3835(25)00745-1

DOI: https://doi.org/10.1016/j.canlet.2025.218173

Reference: CAN 218173

To appear in: *Cancer Letters*

Received Date: 30 June 2025

Revised Date: 19 November 2025

Accepted Date: 24 November 2025

Please cite this article as: A.L. Bowes, S. Waise, T. Lesluyes, T. Butters, C. English, H. Yan, A. Verfaillie, C. Davies, J. Chen, E. Nye, R. Stone, J. Manji, A.M. Flanagan, J. Demeulemeester, M. Tarabichi, N. Pillay, P. Van Loo, Profiling the genomic landscape and evolutionary history of polyploid giant cancer cells in undifferentiated pleomorphic sarcomas, *Cancer Letters*, https://doi.org/10.1016/j.canlet.2025.218173.

Profiling the genomic landscape and evolutionary history of polyploid giant cancer cells in undifferentiated pleomorphic sarcomas

Amy L. Bowes[1,2], Sara Waise[1,3], Tom Lesluyes[1], Thomas Butters[2], Christie English[1], Haixi Yan[1], Annelien Verfaillie[4], Christopher Davies[2], Jianan Chen[2], Emma Nye[5], Richard Stone[5], Jiten Manji[6], Adrienne M. Flanagan[7,11], Jonas Demeulemeester[1,8,9], Maxime Tarabichi[1,10,#], Nischalan Pillay[2,11,#] and Peter Van Loo[1,12,13,#,*]

[1]Cancer Genomics Group, The Francis Crick Institute, London, UK; [2]Sarcoma Biology and Genomics Group, UCL Cancer Institute, London, UK; [3]Cancer Sciences, University of Southampton, Southampton, UK; [4]Genomics Core, UZ / KU Leuven, Belgium; [5]Experimental Histopathology Science and Technology Platform, The Francis Crick Institute, London, UK; [6]Microscopy and Imaging Translational Technology Platform, UCL Cancer Institute, London, UK; [7]Genetics and Cell Biology of Sarcoma, UCL Cancer Institute, London, UK; [8]VIB Centre for Cancer Biology, Leuven, Belgium; [9]Department of Oncology, KU Leuven, Belgium; [10]Institute for Interdisciplinary Research, Université Libre de Bruxelles, Brussels, Belgium; [11]The Royal National Orthopaedic Hospital NHS Trust, London, UK; [12]Department of Genetics, The University of Texas MD Anderson Cancer Centre, Houston, TX, USA; [13]Department of Genomic Medicine, The University of Texas MD Anderson Cancer Centre, Houston, TX, USA.

[#]These authors jointly supervised the study

[*]Corresponding author: pvanloo@mdanderson.org

Abstract

Polyploid giant cancer cells (PGCCs), characterised by multinucleation and atypical nuclear morphology, are a common feature of undifferentiated pleomorphic sarcomas. While PGCCs may be a critical substrate for cancer evolution, their formation pathways and genomic consequences remain underexplored.

In this study, we characterise PGCCs in ten pleomorphic sarcomas and use topographic single-cell DNA sequencing (scDNA-seq) to investigate their genomic landscape. We selected PGCCs based on their nuclear morphology, including mononucleated or multinucleated bizarre, misshapen nuclei, and analysed them at single-cell resolution.

Histopathological analysis showed that PGCCs were often randomly distributed throughout the tumour and did not appear in clusters, suggesting that they arise *de novo* rather than through clonal expansion. scDNA-seq revealed that PGCCs originate from the dominant tumour population and exhibit extensive copy number heterogeneity, either due to subsequent or ongoing chromosomal instability. Both clonal and subclonal chromothripsis-like events were identified in PGCCs, indicating that chromothripsis is a key driver of heterogeneity in these cells and is linked to multinucleation rather than mononuclear PGCC formation. FACS-based ploidy analysis of one undifferentiated pleomorphic sarcoma (UPS) revealed a twice whole-genome-duplicated population (6.2n) distinct from the bulk tumour (3.3n). This population contained all clonal, but none of the subclonal chromothripsis-like events observed in PGCCs. Our findings highlight PGCCs as a highly heterogeneous and evolutionarily dynamic component of UPSs. The recurrent chromothripsis-like events observed in PGCCs suggest ongoing genomic reshaping that may drive tumour progression and the poor clinical outcomes observed for these tumours.

Keywords

Polyploid giant cancer cells, whole genome duplication, chromothripsis, topographic single-cell DNA sequencing, copy number aberration, undifferentiated pleomorphic sarcoma.

Highlights

Topographic scDNA-seq of 10 UPS samples reveals the genomic traits of PGCCs.

PGCCs show random tumour distribution, suggesting *de novo* formation.

PGCCs arise from the dominant tumour population and display rich CNA heterogeneity.

PGCCs contain clonal and subclonal chromothripsis-like genomic events.

Chromothripsis-like events in PGCCs associate with multinucleation.

Introduction

Polyploid giant cancer cells (PGCCs) are a morphologically distinct subpopulation of tumour cells characterised by their markedly enlarged size and high ploidy levels. These cells may contain either a single massive nucleus or multiple nuclei, often 10-20 times larger in size when compared with neighbouring normal diploid cells [1]. PGCCs are thought to arise through successive aberrant cell cycle events involving whole genome duplication (WGD), *i.e.* the duplication of a cell's entire chromosome set during a single aberrant mitotic event. WGD can arise through endoreplication (also referred to as endoreduplication), a process whereby DNA is replicated without subsequent mitosis [2]. Endoreplication is often initiated by mitotic errors, including endocycling, mitotic slippage, endomitosis and cytokinesis failure [3-5]. Non-mitotic processes such as cell fusion and cell cannibalism have also been implicated in PGCC formation [6, 7]. Once formed, these atypical cells may enter dormancy [8] but retain the capacity to reinitiate proliferation and generate a diverse progeny, likely contributing to tumour recurrence, metastasis and treatment resistance [9-11].

WGD is observed in a significant fraction of human cancers [12-14] and underlies many of the morphological hallmarks of PGCCs, including nuclear enlargement, pleomorphism and chromatin architectural changes, all of which are used routinely in common histopathological tumour prognostic grading systems [15, 16]. WGD is linked to increased chromosomal instability, aneuploidy, and enhanced tolerance to large-scale genomic rearrangements, facilitating tumour adaptation and evolutionary plasticity [17-19].

Although WGD events have been well-characterised in bulk sequencing studies, these approaches obscure the nature and contribution of rare cell populations such as PGCCs. Recent advances in single-cell DNA sequencing (scDNA-seq) now permit high-resolution interrogation of tumour heterogeneity, allowing us to study the genomic evolution of individual tumour cells at unprecedented depth [20, 21]. PGCCs frequently show evidence of multiple WGD events [5] and likely harbour large-scale copy number alterations not observed in non-polyploid cells. Therefore, innovative methods such as topographic scDNA-seq, a laser capture-based single-cell isolation and DNA sequencing approach, provides a unique opportunity to map the genomic trajectories of PGCCs *in situ* and elucidate their role in tumour progression [22].

Undifferentiated pleomorphic sarcoma (UPS), a genetically complex subtype of soft tissue sarcoma, represents a powerful model for studying PGCC biology. Up to 90% of UPSs show evidence of at least one WGD and many exhibit multiple duplications (Steele et al., 2019). In

addition, these tumours are heavily enriched for other macroevolutionary genomic events such as chromothripsis, a chromosome shattering and reassembly event associated with poor prognosis in cancer [23-26]. The widespread occurrence of WGD in UPS leads to extensive aneuploidy, genome-wide loss of heterozygosity and complex structural rearrangements. Chromothripsis, often co-occurring with WGD, also permits rapid oncogenic transformation through simultaneous amplification of oncogenes and deletion of tumour suppressors [27].

In this study, we applied topographic scDNA-seq [22] to ten pleomorphic sarcomas to investigate the emergence and clonal architecture of PGCCs. We profiled the genomic landscape of PGCCs in pleomorphic sarcomas to elucidate their clonal relationships and determined how this rare cell type might arise and persist amongst genomically unstable tumour clones.

Results

The spatial distribution and histopathological characteristics of PGCCs

We selected ten pleomorphic sarcomas (French FNCLCC grade 3) for analysis based on current diagnostic criteria (**Table 1**) [28]. The cohort included 7 UPSs, one myxofibrosarcoma, one dedifferentiated liposarcoma (DDLPS) and one pleomorphic liposarcoma. The latter three entities are known to mimic UPS histologically [29].

Across all ten cases, we observed a striking and consistent presence of PGCCs on light microscopy. These cells were defined by markedly abnormal nuclear morphology, including large, irregularly shaped mononuclear or multinuclear forms (**Figure 1a-e**). No osteoclast-type giant cells were observed histologically; these cells derive from the monocyte/macrophage lineage and typically lack abnormal nuclear morphology such as nuclear pleomorphism. PGCCs were typically observed as a randomly dispersed minority population within tumour sections rather than forming discrete clusters. This spatial distribution does not preclude their derivation from a common progenitor but suggests that PGCCs arise *de novo* in parallel while likely retaining their clonal relatedness to the bulk tumour population (**Figure 1a-e**).

Topographic scDNA-seq is an effective method for exploring the genomic landscape of polyploid giant cancer cells in pleomorphic sarcomas

Laser capture microdissection (LCM) permits the precise isolation of individual cells from tissue sections while preserving both their morphological and spatial context. To assess the

potential damaging effect of the laser beam on DNA from PGCC and to exclude introduction of unwanted sequencing artefacts during LCM, we decided to compare the sequencing quality of FACS-sorted single cells to LCM-isolated PGCCs. A total of 20 single cells (11 FACS-sorted nuclei and 9 LCM-isolated PGCCs) underwent WGA and scDNA-seq (**Figure 2a-b**). Notably, LCM-isolated cells exclusively comprised PGCCs, while FACS-sorted nuclei included both normal and aberrant tumour nuclei (**Figure 2a-e**). Comparative analysis revealed that LCM-isolated cells were of high quality relative to FACS-sorted cells. Indeed, LCM-isolated cells exhibited significantly higher percentages of mapped and paired reads when compared with FACS-sorted nuclei (**Figure 2f**), while there was no difference in duplicated reads (**Figure 2g**). They also resulted in superior sequencing coverage breadth (**Figure 2h**), with no observed difference in coverage uniformity assessed by the Gini index, a measure of read distribution evenness (**Figure 2i**). Moreover, copy number calling successfully inferred CNAs in 91% of LCM-isolated cells, compared to 78% of FACS-sorted nuclei (**Figure 2j; Material and Methods**). These findings confirm that LCM is a reliable method for isolating PGCCs for scDNA-seq, despite technical challenges such as navigating polyploid cells through multiple tissue sections.

Polyploid giant cancer cells display extreme copy number heterogeneity

We next performed topographic scDNA-seq on 112 PGCC nuclei from ten pleomorphic sarcomas, leveraging LCM of PGCCs prior to DNA sequencing, preserving their morphological characteristics (**Figure 3a**). PGCC nucleus area ranged from 256-3,080 μm^2 across all ten cases (**Supplementary Figure 1a**). Additional morphometric features of PGCC nuclei are summarised in **Supplementary Figure 1b-f**. To facilitate comparison with the clonal bulk copy number profiles (**Supplementary Figures 2, 3 and 4**), we fitted the sequencing data from PGCCs to a ploidy range closely matching the bulk genome-wide total copy number. We could infer total and allele-specific CNAs in most PGCCs across all ten pleomorphic sarcomas (96 out of 112 cells) that successfully underwent scDNA-seq. PGCC scDNA-seq quality control metrics are shown in **Supplementary Figures 5a-f**. Notably, allele-specific CNAs in PGCCs closely resembled those identified through bulk WGS, suggesting that PGCCs share common genetic attributes with the overall tumour population (**Figure 3b-d**). However, PGCCs also exhibited extreme copy number heterogeneity, with multiple additional copy number gains and losses compared to the bulk tumour population (**Figure 3c-d**) and a high

percentage of the genome being altered (**Figure 3e**), indicating recent, ongoing, or subsequent chromosomal instability.

Polyploid giant cancer cells arise from the clonal tumour population

To investigate genetic relationships among PGCCs and their connection to the dominant clone, we inferred phylogenies of PGCCs isolated from the UPSs using MEDICC2, an allele-specific CNA-based minimum-event distance approach [30]. This revealed extensive CNA heterogeneity within PGCCs, yet all could be traced back to the clonal profile (**Figure 4a-d**). This suggests that PGCCs may emerge *de novo* within tumours as spatially dispersed cells yet still originate from the dominant tumour population before diverging to acquire distinct genetic features, including additional copy number gains and losses. Given the extensive and heterogeneous nature of CNAs observed in the PGCCs, it remains unclear whether these events are under selection. Similar phylogenetic patterns were observed for PGCCs isolated from all ten pleomorphic sarcomas (**Supplementary Figures 6-9**).

Although there was a high degree of copy number heterogeneity, some cases showed consistent (clonal) genomic alterations across all PGCCs isolated from the same tumour (**Figure 4e-f**). This suggests that the observed heterogeneity is not simply due to scDNA-seq noise. For instance, in case PD26861, a small deletion on chromosome 4 was found in every PGCC isolated from that tumour as well as in the bulk-sequenced sample, demonstrating that topographic scDNA-seq can reliably identify small clonal genomic events (**Figure 4f**).

Chromothripsis is a major source of heterogeneity across polyploid giant cancer cells

Chromothripsis is one of the more extensively studied complex rearrangement processes, is prevalent in sarcomas and UPSs, and confers a worse prognosis in multiple human cancers [23-25]. Here, we aimed to explore its prevalence in PGCCs.

Using CTLPScanner, a bioinformatics tool designed to detect chromothripsis-like events, we analysed bulk WGS and scDNA-seq, building on prior studies that demonstrated the applicability of bulk-oriented chromothripsis detection methods to scDNA-seq [31, 32]. Notably, cases PD26882 and PD31205 did not exhibit any clonal chromothripsis-like events. This was consistent with results from two alternative bioinformatic tools, CTCallR and ShatterSeek (**Table 2**, **Material and methods**). Assuming chromothripsis-like events identified in bulk WGS to be clonal, we then leveraged CTLPScanner to screen for these clonal events as well as additional *de novo* chromothripsis-like events in PGCCs.

In the eight pleomorphic sarcomas with clonal chromothripsis calls, CTLPScanner identified at least one or more clonal chromothripsis-like events in 20–83% of PGCCs isolated (**Figure 5a**). In PGCCs where no clonal chromothripsis-like events were identified, it is likely that the resolution of scDNA-seq was insufficient to detect these events. Interestingly, CTLPScanner also detected *de novo* subclonal chromothripsis-like events in a subset of PGCCs (12–90%; **Figure 5b**). Furthermore, *de novo* subclonal chromothripsis-like events were also inferred in PGCCs isolated from tumours lacking clonal chromothripsis events, suggesting that chromothripsis contributes to heterogeneity across PGCCs (**Figure 5b**). The genomic features of clonal (**Figure 5c-d**) and subclonal (**Figure 5e**) chromothripsis-like events were clearly visible in individual PGCCs isolated from the same tumour, although there was some evidence of subclonal diversification (**Figure 5d**). As a negative control, no chromothripsis-like events were identified in single normal diploid cells isolated via FACS (**Figure 5f**).

Whole genome doubled intermediates lack subclonal chromothripsis-like events identified in polyploid giant cancer cells

Because non-PGCC tumour cells cannot be reliably isolated from tissue sections using laser capture microdissection, we employed FACS to separate distinct ploidy subpopulations and compared their CNAs and chromothripsis-like events with PGCCs. FACS-based ploidy analysis in one UPS sample (PD26861) identified a twice WGD population (6.2n) that was distinct from the bulk tumour population (3.3n; **Figure 6a**). This intermediate population carried all clonal chromothripsis events identified in the bulk tumour but lacked the subclonal chromothripsis-like events restricted to PGCCs (**Figure 6b and 6c, Table 3**). Phylogenetic analysis also placed this intermediate twice WGD population on an earlier evolutionary branch, preceding PGCC emergence (**Figure 6d**).

Polyploid giant cancer cell morphology correlates with genomic alterations

Next, we assessed microscopic features such as multinucleation and chromosomal bridging in PGCCs with and without chromothripsis-like events, both clonal and subclonal (**Figure 7a-c**). This analysis revealed a significant difference in multinucleation in PGCCs exhibiting chromothripsis. Specifically, cells with chromothripsis-like events were more likely to be multinucleated PGCCs compared to those without chromothripsis, which often displayed a mononucleated morphology (**Figure 7d; $P = 0.01$**). Lastly, we observed no difference in the

224 rates of chromosome bridge formation between cells with and without chromothripsis-like
225 events (**Figure 7e**).

Discussion

In this study, we examined the evolutionary history and genomic landscape of PGCCs in undifferentiated sarcomas using topographic scDNA-seq. As the isolation of PGCCs is technically challenging due to their increased nuclear content, topographic scDNA-seq is an innovative approach that captures PGCCs from intact tissue sections, permitting the preservation of morphological features.

Our results reveal a heterogeneous genomic landscape and distinct evolutionary trajectories compared to the broader tumour population. Given the frequent presence of PGCCs in UPS compared to carcinoma, their genomes may offer critical insights into sarcoma biology. Unlike many carcinomas, which often arise from identifiable precursor lesions, sarcomas typically lack such defined origins, complicating efforts to pinpoint the early changes that lead to sarcoma formation. Therefore, by studying the genomic characteristics of PGCCs, key pathways driving their genomic evolution may be revealed.

We uncovered extensive genetic heterogeneity amongst PGCCs, with multiple additional CNAs compared to the bulk tumour population. Allele-specific copy number analyses and phylogenetic reconstructions show that PGCCs acquire extra CNAs, likely due to ongoing chromosomal instability. PGCCs thus provide a rich heterogeneous substrate for evolution which may enhance cancer survival and drive tumour progression. In line with this, previous studies have shown that PGCCs exhibit increased migratory and metastatic potential [33, 34]. Although PGCCs appear *de novo* as a spatially dispersed minority population of cells, the presence of shared genomic features indicates a common ancestral origin shared with the main tumour clone. Nonetheless, we acknowledge that while PGCCs can act as a source of genomic diversity, many may also represent evolutionary dead ends. These two possibilities are not mutually exclusive, for example, in some contexts, PGCCs may contribute to clonal diversification and tumour progression, whereas others may undergo senescence or fail to expand clonally. PGCCs may therefore represent a heterogeneous population with the potential to both fuel evolutionary leaps and to terminate without consequence. Although further functional studies will be required to clarify the balance between these two outcomes, we hypothesise PGCCs may only very rarely survive and thrive. When they do, however, they may represent big evolutionary leaps.

Similar CNA patterns were observed in PGCCs from multiple types of high-grade pleomorphic sarcomas, including myxofibrosarcomas, dedifferentiated liposarcomas and pleomorphic

liposarcomas, which can be considered histological mimics of UPS, particularly in small biopsy samples. This suggests a broader relevance of polyploidy and chromothripsis to sarcoma biology beyond UPS. To enable comparison with clonal bulk copy number profiles, we fitted PGCC scDNA-seq data within a ploidy range closely aligned with the bulk genome-wide total copy number. While the actual ploidy of PGCCs is likely substantially higher than the average bulk ploidy, accurately modelling these elevated ploidy levels at single-cell resolution remains logistically and computationally challenging.

Screening for chromothripsis-like events in PGCCs confirmed the presence of clonal chromothripsis events detected in bulk WGS, along with subclonal *de novo* events identified in individual cells. Notably, these *de novo* events were observed even in tumours lacking clonal chromothripsis in bulk sequencing, underscoring the role of chromothripsis in fuelling heterogeneity in PGCCs derived from UPSs. Multinucleated PGCCs were also more likely to harbour chromothripsis-like events than their mononucleated counterparts. Together with our observation that PGCCs sometimes arise in the context of intermediate twice genome doubled populations, which already contain clonal chromothripsis events, but not the subclonal events restricted to PGCCs, these findings offer support for a stepwise trajectory of PGCC evolution. In this model, WGD intermediates may serve as a transitional state and substrate for PGCC formation, with subsequent ploidy increases and the acquisition of subclonal chromothripsis-like events driving genomic diversification and the heterogeneity characteristic of PGCCs. **Figure 8** summarises the possible role of PGCCs in tumour evolution.

Consistent with these findings, prior studies have also shown that PGCCs undergo extensive karyotypic alterations, with spectral karyotyping revealing widespread chromosomal aberrations in their progeny [35]. These findings place chromothripsis within a broader context of stochastic chromosomal alterations, often described as ‘genome chaos’, which fuel tumour evolution by generating bursts of genetic diversity [36]. These chaotic transitions could enable macroevolutionary leaps that reshape the genome, potentially later followed by microevolutionary forces that stabilise advantageous variants and drive tumour adaptation in some PGCCs [37-40]. Novel, high-depth single-cell sequencing approaches are required to resolve these complex karyotypes with greater precision.

Beyond genomic instability, an emerging paradigm in PGCC biology is their capacity to undergo embryonic-like reactivation, a phenomenon substantiated by multiple studies [35, 41]. For example, PGCCs demonstrate molecular and phenotypic hallmarks reminiscent of early embryogenesis, including the induction of pluripotency-associated transcriptional networks,

activation of developmental signalling cascades (e.g., Wnt, Notch, and Hedgehog) and re-expression of stemness markers [41, 42]. This embryonic reprogramming is thought to confer enhanced plasticity, enabling PGCCs to remodel extracellular and stromal architecture, generate heterogeneous cellular progenies through asymmetric division and sustain tumorigenic potential under metabolic or genotoxic stress. The link between these embryonal-like phenotypes and the genomic heterogeneity observed in PGCCs remains to be explored.

In summary, this study provides the first in-depth exploration of the genomes and evolutionary history of PGCCs in pleomorphic sarcoma using topographic scDNA-seq. We reveal that PGCCs frequently exhibit additional genetic alterations, particularly in the form of CNAs and chromothripsis-like events. While PGCCs could be an evolutionary dead end, they may also provide a means for large evolutionary leaps through increased chromosomal instability and chromothripsis. These changes can affect gene dosage, offering a survival advantage to some "hopeful monster" cells, which may contribute to tumour progression and relapse. However, in most PGCCs, these large-scale alterations likely reduce overall fitness. Taken together, this research sheds light on the complex role of PGCCs in cancer progression and their potential implications for tumour evolutionary dynamics.

Materials and Methods

Patient sample acquisition and data collection

Patient tissues and clinical data were collected by the Royal National Orthopaedic Hospital biobank, London Sarcoma Service databases and pathology archives. Patient samples were obtained from the Stanmore Musculoskeletal Biobank, a satellite of the UCL/UCLH Biobank (HTA Licence Number 12055), which was approved by the National Research Ethics Committee (reference 20/YH/0088). This study was approved by the NHS Health Research Authority (REC reference 16/NW/0769).

The pathology archives were searched for sarcomas classified using the international classification of diseases (ICD) codes (http://www.who.int/classifications/icd/en/), coded as undifferentiated sarcoma, pleomorphic sarcoma or spindle cell sarcoma NOS. Only cases with frozen tissue and matching germline material were included. A total of 61 cases were identified. Following pathology review by two pathologists (N.P. and A.B.), 10 high grade sarcomas containing prominent polyploid giant cancer cells and were selected for topographic scDNA-seq. Although a high proportion of PGCC-containing samples in this study are female, this is likely attributable to random variation.

Bulk WGS

Bulk WGS was performed on patient samples using the XTen instrument (Illumina) according to the manufacturer's instructions using 150 bp paired-end libraries. The average coverage was at least 70 X for tumours and 30 X for normal DNA. Bulk UPS WGS was previously published [26].

Laser capture microdissection and image analysis of polyploid giant cancer cells

For LCM, the same snap frozen tumour tissue used for bulk WGS was divided into 5 mm tissue blocks and embedded in Tissue-Tek OCT compound for sectioning. 18 µm thick frozen tissue sections were placed on UV-treated PEN-membrane slides using the Leica CM1950 cryostat at -23°c to -27°c. PEN mounted tissue sections were then fixed in ice cold acetone for 20 minutes and air dried. Tissue sections were then stained with 1% Toluidine blue (Tol blue) for 3 minutes, followed by two washes in DEPC water (15 seconds each) and an incubation in 75%

ethanol for 3 minutes. Once air dried, all tissue sections were imaged using a Zeiss Axio (Z1) slide scanner. Reference tissue sections were stained with H&E.

PGCCs were identified in UPS tissue sections through histological analysis (*i.e.* nuclear size and shape, as well as multinucleation). Single PGCCs were then isolated from fresh frozen tissue sections using a Zeiss PALM Microbeam LCM microscope. Approximately 10 PGCCs per tumour were randomly selected from a small number of serial sections, typically within the same region of the tumour. Several serial sections were taken from a single tumour region and mounted on a PEN-membrane slide. PGCCs were typically tracked through two to four serial sections. Brightfield images were collected before and after the capture of each individual cell, as well as imaging single cells in their corresponding PCR tube following laser powered catapulting. Brightfield images of individual PGCCs were analysed using QuPath (V.0.4.0). Nuclei were manually annotated, and quantitative measurements were obtained for nuclear area, maximum dimension, perimeter, mean optical density and circularity. Among the extracted shape descriptors, solidarity was also calculated, defined as the ratio of the nuclear area to the area of its convex hull. This metric reflects how compact and convex a nucleus is, with values approaching 1 indicating a smooth, solid morphology, and lower values corresponding to irregularity or nuclear indentation.

Whole genome amplification of single cells

Following single-cell isolation by LCM, DNA was extracted and amplified using the Ampli1 WGA kit (BioMenarini Silicon Biosystems) according to the manufacturer's instructions. High sensitivity DNA analysis using a 2100 bioanalyzer (Agilent Technologies) was used to determine the fragment size of amplified single-cell genomic DNA. The ideal library size is between 300-400 bp. The Qubit® double-stranded broad range (dsDNA BR) assay kit (Thermo Fisher Scientific) was used to quantify dsDNA concentrations following single-cell WGA, in the range of 0.1-100 ng/µl, ensuring sufficient input material prior to single-cell library generation and scDNA-seq. In total, 147 polyploid giant cancer cells from 10 pleomorphic sarcoma cases were isolated by LCM and subjected to WGA. 119 single cells underwent successful WGA and scDNA-seq was successfully performed on 112 single cells.

Single-cell DNA library selection and scDNA-seq

Adapters were removed and libraries were size selected using SPRIselect magnetic beads (Beckman Coulter) according to the manufacturer's instructions. In brief, 90 µl of SPRIselect beads were incubated with 50 µl of single-cell DNA (1.8X volume) for 10 minutes at room temperature. Samples were then placed on a magnetic rack for 10 minutes after which the supernatant was discarded. The beads were washed with 200 µl fresh 85% ethanol then air dried on the magnetic rack for 10 minutes. SPRIselect beads were then re-suspended in 40 µl of TE buffer (pH 8.0), mixed and incubated for 10 minutes. Finally, samples were placed on the magnetic rack for a further 10 minutes, after which 30 µl of supernatant was transferred to a fresh tube.

Genomic libraries were prepared using the NEBNext UltraII FS DNA Library Prep Kit for Illumina (New England Biolabs) according to the manufacturer's protocol. Thereafter, DNA was sent for scDNA-seq (Illumina HiSeq 4000) using 150 bp paired-end sequencing with a coverage ranging from 2.5 - 4.5 X per single-cell.

FACS analysis of ploidy and direct library preparation of single nuclei in one UPS

Eight 50-µm frozen tumour sections were dissociated using a detergent-based protocol with a Kimble Dounce tissue grinder. The resulting nuclear suspension was filtered twice through a 40-µm mesh and supplemented with HALT Protease Inhibitor Cocktail (Thermo Scientific, 10320015). Nuclei were stained with 20 µL propidium iodide (PI, 1.0 mg/mL) for at least 40 minutes, then sorted by FACS on a BD FACSAria Fusion cytometer (BD Biosciences, DIVA v9.4). PI-positive nuclei (488 nm excitation, 695/40 filter) were collected, while debris and doublets were excluded. Three populations were gated: a normal diploid control, the bulk tumour population (3.3n), and a twice-WGD tumour population (6.6n). Subpopulations were collected into 384-well plates using the cellenONE system (Cellenion) and processed with the DLP+ protocol [45] using the Nextera i5/i7 library preparation system. Single-cell sequencing libraries were sequenced to a mean depth of ~0.04× per nucleus (Illumina, Nova Xplus).

Single cell WGS alignment and data preprocessing

Paired-end reads generated from bulk WGS and scDNA-seq were aligned to the reference human genome (GRCh38) with BWA-MEM (v0.7.18) [43]. Duplicate reads were marked using MarkDuplicates (Picard v2.23.8) [Broad 44]. SAMtools (v 0.7.17) was used to sort and index aligned reads [45]. SAMtools stats was applied to generate summary statistics for single-

cell BAM files (https://www.htslib.org/). These summary statistics included the total number of sequencing reads, the number of mapped and paired reads, the number of unmapped reads and duplicated reads, the mapping quality distribution, the coverage breadth and the GC content distribution. For the analysis of the FACS WGD tumour subpopulation (6.2n; PD26861), a combined pseudobulk BAM from all single cells was down sampled to match the read depth of single PGCC BAMs.

Bulk WGS bioinformatic analysis

For bulk WGS analysis, somatic variant calling (SNVs, CNAs and SVs) was performed using a validated suite of software with bespoke post-processing filters, which was previously published by Steele *et al*., (2019) [26]. Briefly, Manta (v1.6.0) was used to infer structural variants from bulk WGS data [46]. Co-occurring structural variants and CNAs were visualised across one or multiple chromosomes using an R package called ReConPlot (REarrangement and Copy Number PLOT) [47]. Types of structural variants were annotated as inter-chromosomal translocations, inversions, deletions, or tandem duplications. Bulk chromothripsis calls were generated using CTCallR [27], ShatterSeek [48] and CTLPScanner [49]. The Battenberg algorithm (v2.2.9) was used to infer bulk subclonal copy-number profiles for 10 high grade sarcomas [50].

Single cell bioinformatic analysis

ASCAT.sc was used to infer total and allele-specific copy number aberrations (CNAs) in scDNA-seq data (v0.1; https://github.com/VanLoo-lab/ASCAT.sc). First, ASCAT.sc derives read counts for pre-computed variable sized bins of 30,000 bases for GRCh38 with a mapping quality >= 30, excluding duplicate reads. Second, the log read counts are smoothed by applying a LOESS fit against GC-content to obtain a corrected LogR. Third, a grid search is performed on all ploidy values between 1.0 and 5.5 by steps of 0.01 and a purity of 0.5 or 1 to fit integer copy number values from the LogR track. As ASCAT.sc prefers lower ploidy solutions for some cells, the copy number solutions for some PGCCs were manually refitted to higher ploidy states in agreement with the bulk WGS data. Lastly, a single cell is filtered out when all purity and ploidy solutions result in homozygous deletions of more than 25 Mb, a threshold found to be reached in low quality samples at the bulk level. Allele-specific copy number profiles were derived by leveraging BAF values inferred from matching bulk phased haplotypes derived by the Battenberg algorithm [50].

For single cell phylogenetic analysis, multi-PCF was applied to allele-specific CNAs permitting the segmentation of shared genomic segments with different copy number states across all cells. This produced single-cell copy number profiles with the same set of breakpoints, the outputs of which were then used for downstream phylogenetic tree reconstruction using MEDICC2 (v0.5b) [30]. Alternatively, for the analysis of chromothripsis-like events in single cells, multi-PCF was not employed. This enhanced the sensitivity for detecting clusters of chromosomal breakpoints associated with chromothripsis-like events in scDNA-seq data using CTLPScanner [49]. The percentage of genome altered (PGA) was calculated as the fraction of the genome spanned by segments with total copy number not equal to two, weighted by segment length.

Statistics

No statistical method was used to predetermine cohort size as cases were included based on the patient material and clinical data available. All statistical analyses were performed using GraphPad Prism (v10.2.1). The observed frequencies of morphological characteristics between single cells with and without chromothripsis were assessed using a Fisher's exact test. Significance was set a $P < 0.05$ for all tests.

Tables and Figures

Table 1: Overview of sarcoma cases and their clinicopathological features

Sample ID	Gender	Age at diagnosis	Grade	Nature	Site	Size (mm)	Resection status	Pathological stage	Metastasis	Diagnosis	Survival status
PD26859	F	82	3	Primary	Right shoulder	120	Incomplete	pT2b	Unknown	UPS	Dead
PD26861	F	76	3	Primary	Left thigh	200	Complete	pT2b	Unknown	UPS	Dead
PD26865	F	75	3	Primary	Right thigh	150	Incomplete	pT2b	Soft tissue	UPS	Dead
PD26882	F	60	3	Primary	Left pelvis	230	Complete	pT2b	No	UPS	Alive
PD26890	M	72	3	Primary	Right thigh	250	Marginal	pT2b	Lung and soft tissue	UPS	Dead
PD26902	F	77	3	Recurrence	Right upper arm	15	Marginal	pT2b	No	UPS	Alive
PD26914	F	77	3	Primary	Unknown	85	Complete	pT2b	No	DDLPS	Alive
PD31195	F	40	3	Primary	Unknown	160	Marginal	pT2b	Bone	Pleomorphic liposarcoma	Alive
PD31204	F	69	3	Primary	Right thigh	135	Complete	pT2b	No	Myxofibrosarcoma	Alive
PD31205	M	64	3	Primary	Left thigh	120	Marginal	pT2b	Lung	UPS	Dead

DDLPS = Dedifferentiated liposarcoma. UPS = Undifferentiated pleomorphic sarcoma.

Table 2: Chromothripsis events identified on bulk sequencing.

Cases	Chromothripsis events identified by CTLPScanner	Chromothripsis events identified by ShatterSeek	Chromothripsis events identified by CTCallR
PD26859	chr15, ch18, chr21	None	None
PD26861	chr3, chr4, chr11, chr12, chr18, chr19	chr12	chr12
PD26865	chr1, chr2, chr5, chr6, chr8, chr15	chr1	chr2, chr7, chr8, chr14, chr18
PD26882	None	None	None
PD26890	chr8, chr11	chr9	None
PD26902	chr2, chr5, chr4, chr8, chr11, chr17, chr19, chr20	chr1, chr4, chr8, chr17, chr20	chr4, chr8
PD26914	chr12	None	None
PD31195	chr2, chr10, chr14, chr18	chr3, chr14	None
PD31204	chr2, chr3	chr2	None
PD31205	None	None	None

Table 3: Chromothripsis-like events in PD26861 subpopulations, including the bulk tumour (3.3n), twice WGD subpopulation (6.2n) and single PGCCs, identified by CTLPScanner.

Sample ID	CTLP Region (Mb)	Chromosome	Start	Stop	CNA status change times	Likelihood ratio (log10)
PD26861	83.25	3	1	83247441	39	16
PD26861	46.7	4	125000001	171699983	22	8
PD26861	30	11	5000001	35000000	15	7
PD26861	30	12	100000001	130000000	18	10
PD26861	80.26	18	1	80259271	27	7
PD26861	30	19	28607616	58607616	18	10
2 x WGD	64.33	3	30000001	94334167	33	13
2 x WGD	46.7	4	125000001	171699983	23	7
2 x WGD	30	11	5000001	35000000	15	5
2 x WGD	133.27	12	1	133265309	54	15
2 x WGD	80.26	18	1	80259271	28	6
2 x WGD	30	19	28607616	58607616	17	6
PGCC 2	46.7	4	125000001	171699983	29	8
PGCC 2	133.27	12	1	133265309	58	10
PGCC 2	30	19	28607616	58607616	17	5
PGCC 3	242.18	2	1	242183529	171	10
PGCC 5	83.25	3	5000001	88247441	48	17
PGCC 5	46.7	4	125000001	171699983	21	5
PGCC 5	30	6	85000001	115000000	15	5
PGCC 5	30	11	5000001	35000000	15	5
PGCC 5	30	19	28607616	58607616	15	5
PGCC 7	50.81	14	50000001	100806138	68	5
PGCC 13	58.61	3	30000001	88607616	31	7
PGCC 13	46.7	4	125000001	171699983	26	6
PGCC 13	30	11	5000001	35000000	19	6
PGCC 13	30	12	95000001	125000000	31	16
PGCC 15	30	12	95000001	125000000	24	6
PGCC 16	40	11	5000001	45000000	38	6
PGCC 20	80.26	3	1	80259271	41	9
PGCC 20	30	12	100000001	130000000	24	10

Figure 1: Spatial distribution and histopathological characteristics of PGCCs in pleomorphic sarcomas. (a-b) Representative H&E-stained frozen tissue sections from a pleomorphic sarcoma case (PD26859), demonstrating that PGCCs (white arrows) are randomly distributed throughout the tumour parenchyma rather than being confined to specific regions. The surrounding tumour cells exhibit a heterogenous morphology, ranging from elongated spindled cells to larger epithelioid-appearing cells. Interspersed among the neoplastic cells are scattered immune infiltrates, predominantly small lymphocytes. Scale bar = 200 μm. (c-e) Higher-magnification images of multinucleated PGCCs (white arrows) within the same tumour, illustrating their distinct morphology characterised by enlarged, multi-lobated, hyperchromatic nuclei with anisonucleosis and severe nuclear pleomorphism. Scale bar = 50 μm.

Figure 2: Comparison of topographic single-cell DNA sequencing (scDNA-seq) via laser capture microdissection (LCM) and fluorescence-activated cell sorting (FACS). (a) Schematic of the PALM microbeam LCM system, which employs a pulsed UV laser (wavelength: 337 nm) directed through the epifluorescence path and focused through an objective lens, enabling precise single-cell isolation directly from tissue sections. (b) Representative workflow of FACS used to isolated single nuclei prior to downstream scDNA-seq. Single nuclei were FACS-sorted according to DAPI staining, without selection for specific cell types. The resulting population contained a mixture of tumour and non-tumour cells. (c-e) Examples of PGCCs (indicated by the white and black arrows) being selectively excised from surrounding tumour tissue suing laser ablation and deposited into individual PCR tubes. Scale = 100 μm. (f-j) Quantitative comparison of sequencing performance metrics between FACS-sorted single nuclei and LCM-captured PGCCs. Parameters assessed include the percentage of mapped and paired rates, duplicated reads, coverage breadth across the genome, coverage uniformity as measured by the Gini index (a measure of read distribution evenness) and the proportion of cells yielding successful copy number profiles analysed by ASCAT.sc. Statistical significance was determined by Mann-Whitney tests, with numeric labels denoting P values (threshold for significance: $P < 0.05$). Only significant comparisons are shown.

Figure 3: Genomic heterogeneity of PGCCs revealed by CNA profiling. (a) Approximately ten PGCCs were micro-dissected from each high-grade sarcoma case using laser capture microdissection (LCM). Tumour types included undifferentiated pleomorphic sarcoma (UPS), dedifferentiated liposarcoma (DDLPS), myxofibrosarcoma (MFS) and pleomorphic liposarcoma (PLS). (b) Subclonal allele-specific CNA landscape of a representative UPS case derived from bulk WGS, reflecting the composite genomic profile of the entire tumour cell population. (c-d) Allele-specific CNA profiles of individual PGCCs isolated from the same UPS, displayed at a genomic bin size of 30 Kb. Yellow lines

represent total copy number, while teal lines indicate the copy number of the minor allele. Comparison with the bulk tumour profile highlights both conserved alterations and PGCC-specific genomic differences, illustrating intra-tumoral heterogeneity. (e) Violin plot showing the percentage of the genome altered (PGA) in PGCCs from each sarcoma subtype. Each violin reflects the distribution of PGCC PGA values within a tumour type, while horizontal lines denote the lower quartile, median, and upper quartile. This visualisation demonstrates variability in genomic alteration burden across PGCCs and between sarcoma subtypes.

Figure 4: Phylogenetic relationships and copy number heterogeneity of PGCCs in pleomorphic sarcomas (PD26859 and PD26861). (a-b) Representative histological images of undifferentiated pleomorphic sarcoma (UPS) tissue, demonstrating numerous PGCCs (white arrows). Scale bar = 50 μm. (c-d) Allele-specific copy number-based phylogenetic trees of single PGCCs isolated by laser capture microdissection (LCM) from two independent sarcoma cases (PD26859 and PD26861). Phylogenies illustrate both genetic divergence among individual PGCCs and evidence of a shared common ancestor when compared to the bulk tumour population. Branch lengths correspond to the number of somatic copy number alterations (CNAs) required to transform one profile into another. The 'diploid sample' represents a theoretical normal diploid ancestor. Numbers on the tree denote PGCC IDs; missing IDs indicate cells where library preparation failed. (e-f) Total copy number heatmaps depicting chromosome-wide gains (red) and losses (blue) across single PGCCs from the same cases. Profiles were generated using multi-PCF segmentation with a genomic bin size of 10 Kb. The black arrow highlights a clonal deletion consistently observed in all PGCCs isolated from case PD26861, suggesting a shared ancestral genomic event.

Figure 5: Clonal and subclonal chromothripsis-like events in PGCCs. (a–b) Frequency of chromothripsis-like events detected in PGCCs across ten pleomorphic sarcomas, stratified as (a) *clonal events* (shared with the bulk tumour population) or (b) *subclonal events* (unique to PGCCs and not observed in the bulk profile). Events were inferred using CTLPScanner, highlighting both the persistence of tumour-wide catastrophic events and the emergence of PGCC-specific genomic rearrangements. (c) Representative example of a clonal chromothripsis event detected in bulk WGS, illustrating a complex pattern of localized chromosomal rearrangements. (d–e) Single-cell allele-specific copy number profiles of individual PGCCs from the same tumour case, showing (d) a clonal chromothripsis-like event, concordant with the bulk tumour, and (e) a subclonal chromothripsis-like event unique to the PGCC. For copy number data, black lines represent the total copy number and grey lines denote the minor allele. Structural rearrangements are annotated: black arches = tail-to-tail inversions, green arches = head-to-head inversions, blue arches = tandem duplications, yellow arches =

deletions. (f) Control allele-specific CNA profile from a single normal diploid cell isolated by FACS from the same sarcoma, confirming baseline chromosomal stability. Yellow = total copy number; teal = minor allele copy number.

Figure 6: Chromothripsis-like events in WGD intermediates compared with the bulk tumour and PGCCs (PD26861). (a) FACS ploidy histogram from a single UPS sample (PD26861) showing a distinct twice WGD (6.2n) population compared with the bulk tumour (3.3n) and a diploid control cell line. (b) Distribution of clonal versus subclonal chromothripsis-like events, where clonal events are shared between the bulk tumour and the twice WGD population, whereas subclonal chromothripsis-like events appear restricted to PGCCs. (c) Heatmap of individual chromothripsis-like events for the bulk tumour (3.3n), the twice WGD population (6.2n) and PGCCs isolated from the same tumour via LCM. (d) Allele-specific copy number-based phylogenetic tree for an UPS, including the bulk tumour, twice WGD intermediate population and single PGCCs (PD26861). Phylogenetic analysis suggests that the twice-WGD intermediate precedes PGCC formation supporting a stepwise trajectory of PGCC evolution. Branch lengths correspond to the number of somatic copy number alterations (CNAs) required to transform one profile into another. The 'diploid sample' represents a theoretical normal diploid ancestor. Numbers on the tree denote PGCC IDs; missing IDs indicate cells where library preparation failed.

Figure 7: Histopathological features of single PGCCs stratified by the presence or absence of chromothripsis-like events. (a–c) Representative histological images of polyploid giant cancer cells (PGCCs, white arrows) captured on H&E-stained frozen tissue sections, illustrating their diverse morphologies: (a) mononucleated PGCC, (b) multinucleated PGCC, and (c) PGCC undergoing chromosomal bridge formation, a feature suggestive of mitotic errors and structural genome instability. Scale bar = 50 μm. (d–e) Quantitative comparison of histopathological features between PGCCs with and without chromothripsis-like events. Fisher's exact test was used to assess significance. (d) Multinucleated PGCCs were significantly enriched for chromothripsis-like events compared with mononucleated PGCCs ($p = 0.01$), suggesting an association between multinucleation and catastrophic genomic rearrangements. (e) In contrast, no significant difference was observed in the frequency of chromosomal bridge formation between PGCCs harbouring or lacking chromothripsis-like events ($p = 0.48$). Non-significant P values are not displayed. Numbers above each bar denote the number of PGCCs analysed in each category.

Figure 8: PGCCs and tumour evolution: (a) Polyploid giant cancer cells (PGCCs) display a random distribution within undifferentiated pleomorphic sarcomas (UPSs), suggesting de novo formation.

Owing to multiple formation pathways, including endocycling, mitotic slippage, endomitosis and cytokinesis failure, PGCCs exhibit diverse morphologies and may appear either mononucleated or multinucleated. (b) Chromothripsis results from a catastrophic cellular event that causes chromosome shattering and random reassembly, leading to extensively rearranged chromosomes, loss of DNA fragments, and the formation of extrachromosomal DNA (ecDNA). (c) Proposed model for PGCC formation. Whole-genome duplication (WGD) intermediates may act as transitional states that give rise to PGCCs. Alternatively, PGCCs can also arise after a catastrophic mitotic failure, such as chromothripsis. Subsequent increases in ploidy and acquisition of subclonal chromothripsis-like events may contribute to genomic diversification and the intra-tumoural heterogeneity characteristic of PGCCs.

CRediT authorship contribution statement

Amy L. Bowes: Writing - review & editing, Writing - original draft, Project administration, Methodology, Formal analysis, Data curation, Conceptualisation.

Sara Waise: Writing - review & editing, Methodology, Formal analysis.

Tom Lesluyes: Writing - review & editing, Methodology, Formal analysis.

Christie English: Methodology.

Haixi Yan: Methodology.

Annelien Verfaillie: Methodology.

Thomas Butters: Methodology, Formal analysis.

Christie Davies: Methodology.

Jianan Chen: Methodology.

Emma Nye: Methodology.

Richard Stone: Methodology.

Jiten Manji: Methodology.

Adrienne M. Flanagan: Methodology.

Jonas Demeulemeester: Methodology, Writing - review & editing.

Maxime Tarabichi: Supervision, Conceptualisation, Writing - review & editing, Methodology, Formal analysis.

Nischalan Pillay: Supervision, Project administration, Funding acquisition, Conceptualisation, Writing - review & editing.

Peter Van Loo: Supervision, Project administration, Funding acquisition, Conceptualisation, Writing - review & editing.

Supplementary data
The following is the supplementary data to this article.

Supplementary Figure 1: Nuclear morphometric features of PGCCs. (a-f) Violin plots illustrating the distribution of nuclear morphologies among PGCCs isolated from each of the ten pleomorphic sarcoma samples. Quantitative features assessed include nuclear area, perimeter, circularity, aspect ratio and solidity. Solidity, defined as the ratio of nuclear area to its convex hull, provides a measure of nuclear compactness and convexity, with values closer to 1 indicating smoother, more circular nuclei. Each violin plot displays the full range of measurements within a sample, while horizontal lines denote the lower quartile, mean and upper quartile. These analyses demonstrate both inter- and intra-tumoral variability in PGCC nuclear architecture, reflecting their morphological heterogeneity.

Supplementary Figure 2: Bulk subclonal allele-specific copy number profiles for cases PD26859, PD26861, PD26865, PD26882 and PD26890. Allele-specific copy number plots are shown for cases PD26859, PD26861, PD26865, PD26882 and PD26890 across the entire genome. For each chromosome, allele-specific copy number states are represented as yellow (major allele) and teal (minor allele) lines. Non-integer copy number values reflect the presence of subclonal tumour populations, with the relative size of each subclone inferred from the distance between adjacent integer states. This approach highlights the coexistence of clonal and subclonal genomic alterations within the same tumour.

Supplementary Figure 3: Bulk subclonal allele-specific copy number profiles for cases PD26902, PD26914, PD31195, PD31205 and PD31204. Allele-specific copy number profiles are shown for cases PD26902, PD26914, PD31195, PD31205 and PD31204 across the entire genome. For each chromosome, allele-specific copy number states are represented as yellow (major allele) and teal (minor allele) lines. Non-integer copy number values reflect the presence of subclonal tumour populations, with the relative size of each subclone inferred from the distance between adjacent integer states. This approach highlights the coexistence of clonal and subclonal genomic alterations within the same tumour.

Supplementary Figure 4: Ploidy optimisation plots for PGCC allele-specific CNAs: (a) Allele-specific copy number profile for bulk tumour sample PD26861. (b) Allele-specific copy number profile for a single PGCC (PD26861.5) isolated from the same case, fitted to a maximum ploidy of 3.5 to facilitate direct visual comparison with the bulk tumour profile. (c) Alternative fitting of the same PGCC sample (PD26861.5) to a maximum ploidy of 4.5, which introduces greater visual complexity. For each

chromosome, allele-specific copy number states are shown as yellow (major allele) and teal (minor allele) lines. Non-integer copy number values indicate the presence of subclonal tumour populations, with the relative size of each subclone inferred from the spacing between adjacent integer states.

Supplementary Figure 5: Quality control metrics for PGCC scDNA-seq. (a) Bar plot summarising the number of PGCCs per case that underwent successful whole-genome amplification (WGA). Numbers above each bar indicate the number of successfully amplified single cells. (b–f) Scatter plots showing individual quality control metrics for each successfully sequenced PGCC, with each dot representing one single cell. Metrics assessed include DNA yield, duplicated sequencing reads, unmapped reads, mapped and paired reads and coverage breadth across the whole genome. Together, these parameters provide an overview of sequencing quality and the reliability of downstream copy number profiling.

Supplementary Figure 6: Phylogenetic relationships and copy number heterogeneity of PGCCs in pleomorphic sarcomas (PD26865 and PD26882). (a-b) Representative histological images of undifferentiated pleomorphic sarcoma (UPS) tissue, demonstrating numerous PGCCs (white arrows). Scale bar = 50 μm. (c-d) Allele-specific copy number-based phylogenetic trees of single PGCCs isolated by laser capture microdissection (LCM) from two independent sarcoma cases (PD26865 and PD26882). Phylogenies illustrate both genetic divergence among individual PGCCs and evidence of a shared common ancestor when compared with the bulk tumour population. Branch length corresponds to the number of somatic copy number alterations (CNAs) required to transform one profile into another. The 'diploid sample' represents a theoretical normal diploid ancestor. Numbers on the tree denote PGCC IDs; missing IDs indicate cells where library preparation failed. (e-f) Total copy number heatmaps depicting chromosome-wide gains (red) and losses (blue) across single PGCCs from the same cases. Profiles were generated using multi-PCF segmentation with a genomic bin size of 10 Kb.

Supplementary Figure 7: Phylogenetic relationships and copy number heterogeneity of PGCCs in pleomorphic sarcomas (PD26890 and PD26902). (a-b) Representative histological images of undifferentiated pleomorphic sarcoma (UPS) tissue, demonstrating numerous PGCCs (white arrows). Scale bar = 50 μm. (c-d) Allele-specific copy number-based phylogenetic trees of single PGCCs isolated by laser capture microdissection (LCM) from two independent sarcoma cases (PD26890 and PD26902). Phylogenies illustrate both genetic divergence among individual PGCCs and evidence of a shared common ancestor when compared with the bulk tumour population. Branch length corresponds to the number of somatic copy number alterations (CNAs) required to transform one profile into another. The 'diploid sample' represents a theoretical normal diploid ancestor. Numbers on the tree denote PGCC IDs; missing IDs indicate cells where library preparation failed. (e-f) Total copy number heatmaps

depicting chromosome-wide gains (red) and losses (blue) across single PGCCs from the same cases. Profiles were generated using multi-PCF segmentation with a genomic bin size of 10 Kb.

Supplementary Figure 8: Phylogenetic relationships and copy number heterogeneity of PGCCs in a dedifferentiated liposarcoma (DDLPS) and pleomorphic liposarcoma (PD26914 and PD31195). (a-b) Representative histological images of a DDLPS and pleomorphic liposarcoma, demonstrating numerous PGCCs (white arrows). Scale bar = 50 μm. (c-d) Allele-specific copy number-based phylogenetic trees of single PGCCs isolated by laser capture microdissection (LCM) from two independent sarcoma cases (PD26914 and PD31195). Phylogenies illustrate both genetic divergence among individual PGCCs and evidence of a shared common ancestor when compared with the bulk tumour population. Branch length corresponds to the number of somatic copy number alterations (CNAs) required to transform one profile into another. The 'diploid sample' represents a theoretical normal diploid ancestor. Numbers on the tree denote PGCC IDs; missing IDs indicate cells where library preparation failed. (e-f) Total copy number heatmaps depicting chromosome-wide gains (red) and losses (blue) across single PGCCs from the same cases. Profiles were generated using multi-PCF segmentation with a genomic bin size of 10 Kb.

Supplementary Figure 9: Phylogenetic relationships and copy number heterogeneity of PGCCs in pleomorphic sarcomas (PD31204 and PD31205). (a-b) Representative histological images of a myxofibrosarcoma and undifferentiated pleomorphic sarcoma (UPS), demonstrating numerous PGCCs (white arrows). Scale bar = 50 μm. (c-d) Allele-specific copy number-based phylogenetic trees of single PGCCs isolated by laser capture microdissection (LCM) from two independent sarcoma cases (PD31204 and PD31205). Phylogenies illustrate both genetic divergence among individual PGCCs and evidence of a shared common ancestor when compared with the bulk tumour population. Branch length corresponds to the number of somatic copy number alterations (CNAs) required to transform one profile into another. The 'diploid sample' represents a theoretical normal diploid ancestor. Numbers on the tree denote PGCC IDs; missing IDs indicate cells where library preparation failed. (e-f) Total copy number heatmaps depicting chromosome-wide gains (red) and losses (blue) across single PGCCs from the same cases. Profiles were generated using multi-PCF segmentation with a genomic bin size of 10 Kb.

Data availability

The data supporting the findings of this study are available within the paper and its supplementary information files. Single cell BAM files are available to download from the European Genome Archive (EGA) database at accension number EGAXXX. Bulk BAM files are available to download from EGA at accension number EGAD00001004162 and were previously published [26].

Funding

AB, TL, SW, MT, and PVL were supported by the Francis Crick Institute, which receives its core funding from Cancer Research UK (CC2008), the UK Medical Research Council (CC2008), and the Wellcome Trust (CC2008). J.D. was supported by a postdoctoral fellowship from the Research Foundation – Flanders (project no. 12J6916N) and acknowledges current grant support from VIB (Vlaams Instituut voor Biotechnologie). MT was a supported postdoctoral researcher of the F.R.S.-FNRS. PVL is a Winton Group Leader in recognition of the Winton Charitable Foundation's support towards the establishment of the Francis Crick Institute. PVL is a CPRIT Scholar in Cancer Research and acknowledges CPRIT grant support (RR210006). NP holds a Cancer Research UK Career Establishment award (Ref: RCCCEA-Nov23/100003), received funding for the sample acquisition and sequencing through a Cancer Research UK Clinician Scientist Fellowship (Grant no: 18387) and is supported by the UCL-Cancer Research UK Experimental Cancer Centre.

Declaration of competing interests

The authors declare no conflicts of interests.

Acknowledgments

The authors thank Christopher Steele for providing matching bulk WGS data and the Wellcome Sanger Institute for performing bulk WGS. We also thank Erik Sahai and his team for their support during this project. Figures were created using Biorender.com.

References

[1] J. Liu, K. Liu, M. Yan, J. Xu, Y. Liu, S. Zhang, Dormant and activated polyploid giant cancer cells: from chemoradiotherapy resistance to cancer progression, Cancer Lett, (2025) 218137.

[2] G. Temaj, S. Saha, S. Chichiarelli, P. Telkoparan-Akillilar, N. Nuhii, R. Hadziselimovic, L. Saso, Polyploid giant cancer cells: Underlying mechanisms, signaling pathways, and therapeutic strategies, Crit Rev Oncol Hematol, 213 (2025) 104802.

[3] R.S. Balachandran, E.T. Kipreos, Addressing a weakness of anticancer therapy with mitosis inhibitors: Mitotic slippage, Mol Cell Oncol, 4 (2017) e1277293.

[4] H. Was, A. Borkowska, A. Olszewska, A. Klemba, M. Marciniak, A. Synowiec, C. Kieda, Polyploidy formation in cancer cells: How a Trojan horse is born, Semin Cancer Biol, 81 (2022) 24-36.

[5] K.J. Pienta, E.U. Hammarlund, R.H. Austin, R. Axelrod, J.S. Brown, S.R. Amend, Cancer cells employ an evolutionarily conserved polyploidization program to resist therapy, Semin Cancer Biol, 81 (2022) 145-159.

[6] Y. Ogawa, L. Fisher, T. Matsumoto, The Impact of Polyploid Giant Cancer Cells: The Root of Stress Resilience, Cancer Sci, 116 (2025) 2949-2958.

[7] D.L. Adams, M. Cristofanilli, S.H. Lin, R.C. Bergan, T.H. Ho, J.R. Marks, S.S. Martin, M.J. Edelman, S. Chumsri, E.J. Hager, C.M. Tang, S. Tsai, R.K. Alpaugh, Phenotyping and clinical utility of phagocytic polyploid giant cancer macrophages in blood, Cancer Lett, 631 (2025) 218007.

[8] Y. Jiao, Y. Yu, M. Zheng, M. Yan, J. Wang, Y. Zhang, S. Zhang, Dormant cancer cells and polyploid giant cancer cells: The roots of cancer recurrence and metastasis, Clin Transl Med, 14 (2024) e1567.

[9] A.N. Makarovskiy, E. Siryaporn, D.C. Hixson, W. Akerley, Survival of docetaxel-resistant prostate cancer cells in vitro depends on phenotype alterations and continuity of drug exposure, Cell Mol Life Sci, 59 (2002) 1198-1211.

[10] K. Mittal, S. Donthamsetty, R. Kaur, C. Yang, M.V. Gupta, M.D. Reid, D.H. Choi, P.C.G. Rida, R. Aneja, Multinucleated polyploidy drives resistance to Docetaxel chemotherapy in prostate cancer, Br J Cancer, 116 (2017) 1186-1194.

[11] A. Ogden, P.C. Rida, B.S. Knudsen, O. Kucuk, R. Aneja, Docetaxel-induced polyploidization may underlie chemoresistance and disease relapse, Cancer Lett, 367 (2015) 89-92.

[12] C.M. Bielski, A. Zehir, A.V. Penson, M.T.A. Donoghue, W. Chatila, J. Armenia, M.T. Chang, A.M. Schram, P. Jonsson, C. Bandlamudi, P. Razavi, G. Iyer, M.E. Robson, Z.K. Stadler, N. Schultz, J. Baselga, D.B. Solit, D.M. Hyman, M.F. Berger, B.S. Taylor, Genome doubling shapes the evolution and prognosis of advanced cancers, Nat Genet, 50 (2018) 1189-1195.

[13] K. Prasad, M. Bloomfield, H. Levi, K. Keuper, S.V. Bernhard, N.C. Baudoin, G. Leor, Y. Eliezer, M. Giam, C.K. Wong, G. Rancati, Z. Storchová, D. Cimini, U. Ben-David, Whole-Genome Duplication Shapes the Aneuploidy Landscape of Human Cancers, Cancer Res, 82 (2022) 1736-1752.

[14] C.D. McKenney, S. Regot, Mechanisms of whole-genome doubling in cancer evolution, Trends Cancer, (2025).

[15] C.W. Elston, I.O. Ellis, Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up, Histopathology, 19 (1991) 403-410.

[16] B. Delahunt, J.N. Eble, L. Egevad, H. Samaratunga, Grading of renal cell carcinoma, Histopathology, 74 (2019) 4-17.

[17] S.M. Dewhurst, N. McGranahan, R.A. Burrell, A.J. Rowan, E. Grönroos, D. Endesfelder, T. Joshi, D. Mouradov, P. Gibbs, R.L. Ward, N.J. Hawkins, Z. Szallasi, O.M. Sieber, C. Swanton, Tolerance of whole-genome doubling propagates chromosomal instability and accelerates cancer genome evolution, Cancer Discov, 4 (2014) 175-185.

[18] T.B.K. Watkins, E.L. Lim, M. Petkovic, S. Elizalde, N.J. Birkbak, G.A. Wilson, D.A. Moore, E. Grönroos, A. Rowan, S.M. Dewhurst, J. Demeulemeester, S.C. Dentro, S. Horswell, L. Au, K. Haase, M. Escudero, R. Rosenthal, M.A. Bakir, H. Xu, K. Litchfield, W.T. Lu, T.P. Mourikis, M. Dietzen, L. Spain, G.D. Cresswell, D. Biswas, P. Lamy, I. Nordentoft, K. Harbst, F. Castro-Giner, L.R. Yates, F. Caramia, F. Jaulin, C. Vicier, I.P.M. Tomlinson, P.K. Brastianos, R.J. Cho, B.C. Bastian, L. Dyrskjøt, G.B. Jönsson, P. Savas, S. Loi, P.J. Campbell, F. Andre, N.M. Luscombe, N. Steeghs, V.C.G. Tjan-Heijnen, Z. Szallasi, S. Turajlic, M. Jamal-Hanjani, P. Van Loo, S.F. Bakhoum, R.F. Schwarz, N. McGranahan, C. Swanton, Pervasive chromosomal instability and karyotype order in tumour evolution, Nature, 587 (2020) 126-132.

[19] H.C. Herriage, C.L. Hughes, S.K. Fahey, B.R. Calvi, Unscheduled polyploidy synergizes with oncogenic mutations to enhance genome instability and tumorigenesis, Cancer Lett, (2025) 218008.

[20] A.L. Bowes, M. Tarabichi, N. Pillay, P. Van Loo, Leveraging single-cell sequencing to unravel intratumour heterogeneity and tumour evolution in human cancers, J Pathol, 257 (2022) 466-478.

[21] B. Lu, Cancer phylogenetic inference using copy number alterations detected from DNA sequencing data, Cancer Pathog Ther, 3 (2025) 16-29.

[22] A.K. Casasent, A. Schalck, R. Gao, E. Sei, A. Long, W. Pangburn, T. Casasent, F. Meric-Bernstam, M.E. Edgerton, N.E. Navin, Multiclonal Invasion in Breast Tumors Identified by Topographic Single Cell Sequencing, Cell, 172 (2018) 205-217.e212.

[23] F. Magrangeas, H. Avet-Loiseau, N.C. Munshi, S. Minvielle, Chromothripsis identifies a rare and aggressive entity among newly diagnosed multiple myeloma patients, Blood, 118 (2011) 675-678.

[24] T. Rausch, D.T. Jones, M. Zapatka, A.M. Stütz, T. Zichner, J. Weischenfeldt, N. Jäger, M. Remke, D. Shih, P.A. Northcott, E. Pfaff, J. Tica, Q. Wang, L. Massimi, H. Witt, S. Bender, S. Pleier, H. Cin, C. Hawkins, C. Beck, A. von Deimling, V. Hans, B. Brors, R. Eils, W. Scheurlen, J. Blake, V. Benes, A.E. Kulozik, O. Witt, D. Martin, C. Zhang, R. Porat, D.M. Merino, J. Wasserman, N. Jabado, A. Fontebasso, L. Bullinger, F.G. Rücker, K. Döhner, H. Döhner, J. Koster, J.J. Molenaar, R. Versteeg, M. Kool, U. Tabori, D. Malkin, A. Korshunov, M.D. Taylor, P. Lichter, S.M. Pfister, J.O. Korbel, Genome sequencing of pediatric medulloblastoma links catastrophic DNA rearrangements with TP53 mutations, Cell, 148 (2012) 59-71.

[25] D. Hirsch, R. Kemmerling, S. Davis, J. Camps, P.S. Meltzer, T. Ried, T. Gaiser, Chromothripsis and focal copy number alterations determine poor outcome in malignant melanoma, Cancer Res, 73 (2013) 1454-1460.

[26] C.D. Steele, M. Tarabichi, D. Oukrif, A.P. Webster, H. Ye, M. Fittall, P. Lombard, I. Martincorena, P.S. Tarpey, G. Collord, K. Haase, S.J. Strauss, F. Berisha, H. Vaikkinen, P. Dhami, M. Jansen, S. Behjati, M.F. Amary, R. Tirabosco, A. Feber, P.J. Campbell, L.B. Alexandrov, P. Van Loo, A.M. Flanagan, N. Pillay, Undifferentiated Sarcomas Develop through Distinct Evolutionary Pathways, Cancer Cell, 35 (2019) 441-456.e448.

[27] The ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium, *Pan-cancer analysis of whole genomes.*, Nature, 578 (2020) 82-93.

[28] WHO, WHO Classification of Tumours of Soft Tissue and Bone, Lyon, France., 2020.

[29] C.D. Fletcher, P. Gustafson, A. Rydholm, H. Willén, M. Akerman, Clinicopathologic re-evaluation of 100 malignant fibrous histiocytomas: prognostic relevance of subclassification, J Clin Oncol, 19 (2001) 3045-3050.

[30] T.L. Kaufmann, M. Petkovic, T.B.K. Watkins, E.C. Colliver, S. Laskina, N. Thapa, D.C. Minussi, N. Navin, C. Swanton, P. Van Loo, K. Haase, M. Tarabichi, R.F. Schwarz, MEDICC2: whole-genome doubling aware copy-number phylogenies for cancer evolution, Genome Biol, 23 (2022) 241.

[31] P. Smirnov, M.J. Przybilla, M. Simovic-Lorenz, R.G. Parra, H. Susak, M. Ratnaparkhe, J.K. Wong, V. Körber, J.P. Mallm, G. Philippos, M. Sill, T. Kolb, R. Kumar, N. Casiraghi, K. Okonechnikov, D.R. Ghasemi, K.K. Maaß, K.W. Pajtler, A. Jauch, A. Korshunov, T. Höfer, M. Zapatka, S.M. Pfister, W. Huber, O. Stegle, A. Ernst, Multi-omic and single-cell profiling of chromothriptic medulloblastoma reveals genomic and transcriptomic consequences of genome instability, Nat Commun, 15 (2024) 10183.

[32] B.R. Mardin, A.P. Drainas, S.M. Waszak, J. Weischenfeldt, M. Isokane, A.M. Stütz, B. Raeder, T. Efthymiopoulos, C. Buccitelli, M. Segura-Wang, P. Northcott, S.M. Pfister, P. Lichter, J. Ellenberg, J.O. Korbel, A cell-based model system links chromothripsis with hyperploidy, Mol Syst Biol, 11 (2015) 828.

[33] M.M. Mallin, N. Kim, M.I. Choudhury, S.J. Lee, S.S. An, S.X. Sun, K. Konstantopoulos, K.J. Pienta, S.R. Amend, Cells in the polyaneuploid cancer cell (PACC) state have increased metastatic potential, Clin Exp Metastasis, 40 (2023) 321-338.

[34] L. Trabzonlu, K.J. Pienta, B.J. Trock, A.M. De Marzo, S.R. Amend, Presence of cells in the polyaneuploid cancer cell (PACC) state predicts the risk of recurrence in prostate cancer, Prostate, 83 (2023) 277-285.

[35] N. Niu, J. Zhang, N. Zhang, I. Mercado-Uribe, F. Tao, Z. Han, S. Pathak, A.S. Multani, J. Kuang, J. Yao, R.C. Bast, A.K. Sood, M.C. Hung, J. Liu, Linking genomic reorganization to tumor initiation via the giant cell cycle, Oncogenesis, 5 (2016) e281.

[36] H.H. Heng, J.B. Stevens, G. Liu, S.W. Bremer, K.J. Ye, P.V. Reddy, G.S. Wu, Y.A. Wang, M.A. Tainsky, C.J. Ye, Stochastic cancer progression driven by non-clonal chromosome aberrations, J Cell Physiol, 208 (2006) 461-472.

[37] J. Liu, The "life code": A theory that unifies the human life cycle and the origin of human tumors, Semin Cancer Biol, 60 (2020) 380-397.

[38] J. Heng, H.H. Heng, Genome chaos: Creating new genomic information essential for cancer macroevolution, Semin Cancer Biol, 81 (2022) 160-175.

[39] X. Li, Y. Zhong, X. Zhang, A.K. Sood, J. Liu, Spatiotemporal view of malignant histogenesis and macroevolution via formation of polyploid giant cancer cells, Oncogene, 42 (2023) 665-678.

[40] X. Zhang, J. Yao, X. Li, N. Niu, Y. Liu, R.A. Hajek, G. Peng, S. Westin, A.K. Sood, J. Liu, Targeting polyploid giant cancer cells potentiates a therapeutic response and overcomes resistance to PARP inhibitors in ovarian cancer, Sci Adv, 9 (2023) eadf7195.

[41] Q. Song, M. Gao, Y. Weng, X. Zhuang, Y. Wu, H. Cui, N. Ding, L. Wang, S. Bi, L. Zhang, W. Zhang, Y. Cui, Evolutionary adaptation and asymmetric inheritance of polyploid giant cancer cells in esophageal squamous cell carcinoma, Cancer Lett, (2025) 217818.

[42] S. Zhang, I. Mercado-Uribe, Z. Xing, B. Sun, J. Kuang, J. Liu, Generation of cancer stem-like cells through the formation of polyploid giant cancer cells, Oncogene, 33 (2014) 116-128.

[43] H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform, Bioinformatics, 25 (2009) 1754-1760.

[44] B. Institute, Picard toolkit, 2019.

[45] P. Danecek, J.K. Bonfield, J. Liddle, J. Marshall, V. Ohan, M.O. Pollard, A. Whitwham, T. Keane, S.A. McCarthy, R.M. Davies, H. Li, Twelve years of SAMtools and BCFtools, Gigascience, 10 (2021).

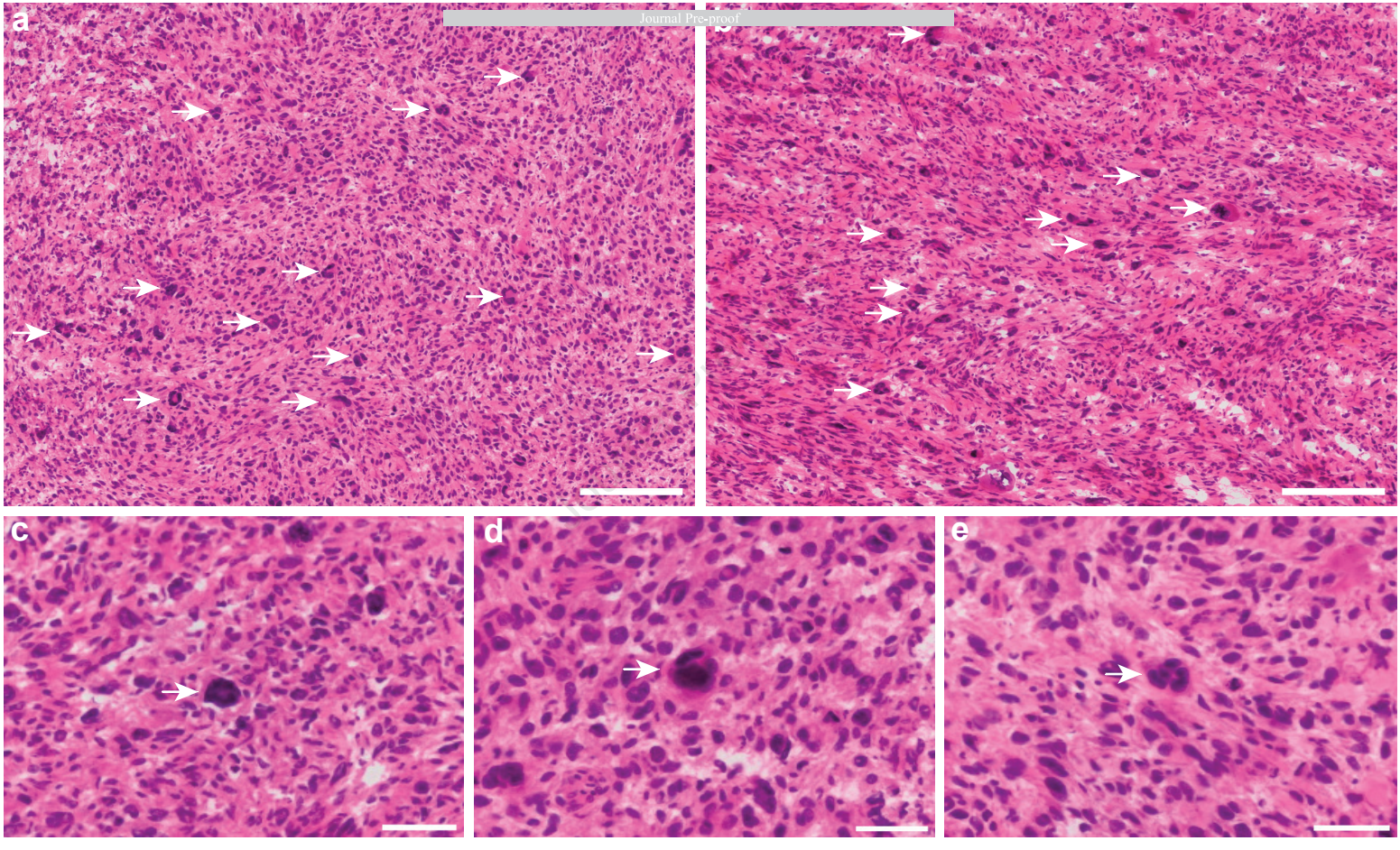
[46] X. Chen, O. Schulz-Trieglaff, R. Shaw, B. Barnes, F. Schlesinger, M. Källberg, A.J. Cox, S. Kruglyak, C.T. Saunders, Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications, Bioinformatics, 32 (2016) 1220-1222.

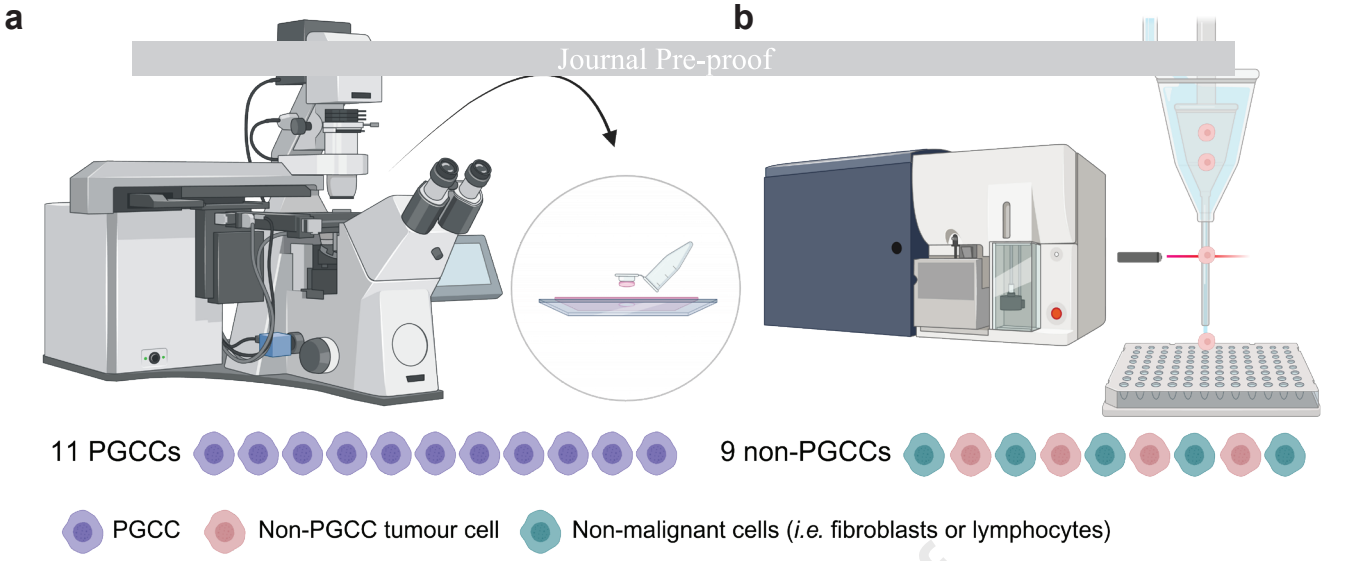
[47] J. Espejo Valle-Inclán, I. Cortés-Ciriano, ReConPlot: an R package for the visualization and interpretation of genomic rearrangements, Bioinformatics, 39 (2023).

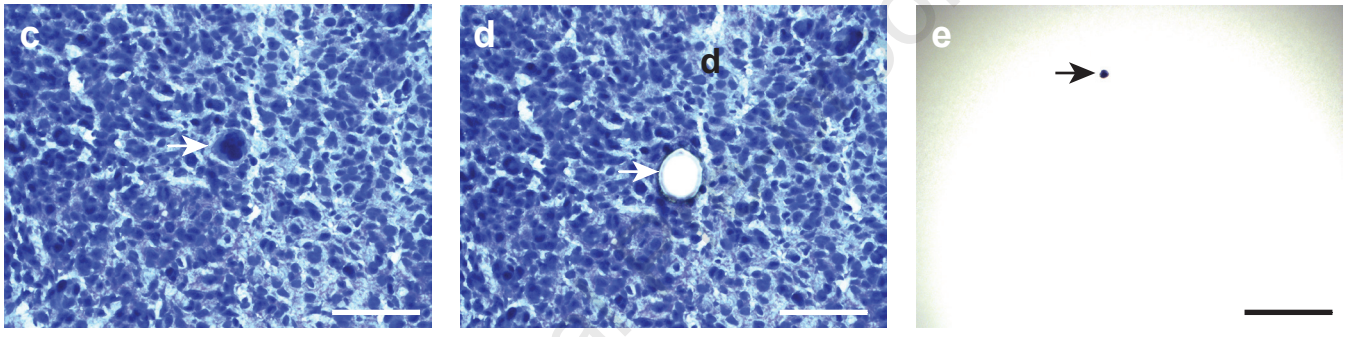
[48] I. Cortés-Ciriano, J.J. Lee, R. Xi, D. Jain, Y.L. Jung, L. Yang, D. Gordenin, L.J. Klimczak, C.Z. Zhang, D.S. Pellman, P.J. Park, Comprehensive analysis of chromothripsis in 2,658 human cancers using whole-genome sequencing, Nat Genet, 52 (2020) 331-341.

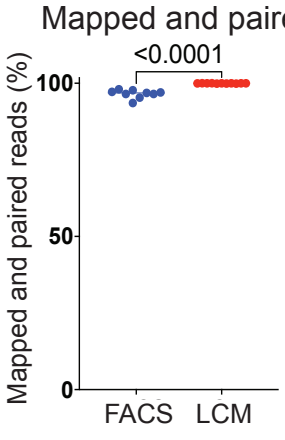
[49] J. Yang, J. Liu, L. Ouyang, Y. Chen, B. Liu, H. Cai, CTLPScanner: a web server for chromothripsis-like pattern detection, Nucleic Acids Res, 44 (2016) W252-258.

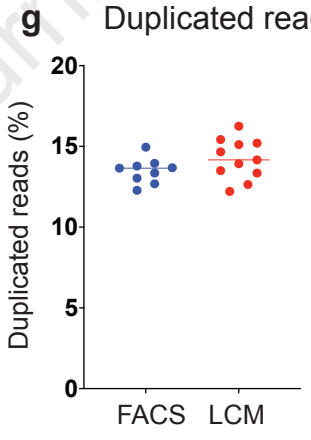
[50] S. Nik-Zainal, P. Van Loo, D.C. Wedge, L.B. Alexandrov, C.D. Greenman, K.W. Lau, K. Raine, D. Jones, J. Marshall, M. Ramakrishna, A. Shlien, S.L. Cooke, J. Hinton, A. Menzies, L.A. Stebbings, C. Leroy, M. Jia, R. Rance, L.J. Mudie, S.J. Gamble, P.J. Stephens, S. McLaren, P.S. Tarpey, E. Papaemmanuil, H.R. Davies, I. Varela, D.J. McBride, G.R. Bignell, K. Leung, A.P. Butler, J.W. Teague, S. Martin, G. Jönsson, O. Mariani, S. Boyault, P. Miron, A. Fatima, A. Langerød, S.A. Aparicio, A. Tutt, A.M. Sieuwerts, Å. Borg, G. Thomas, A.V. Salomon, A.L. Richardson, A.L. Børresen-Dale, P.A. Futreal, M.R. Stratton, P.J. Campbell, The life history of 21 breast cancers, Cell, 149 (2012) 994-1007.

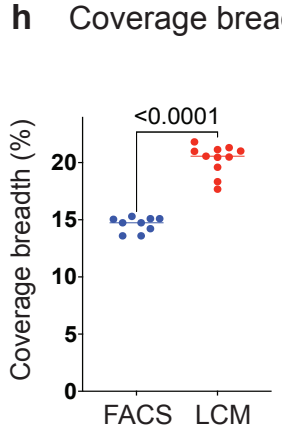


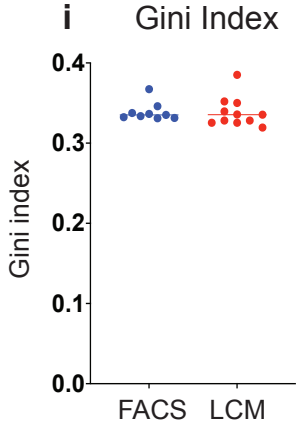


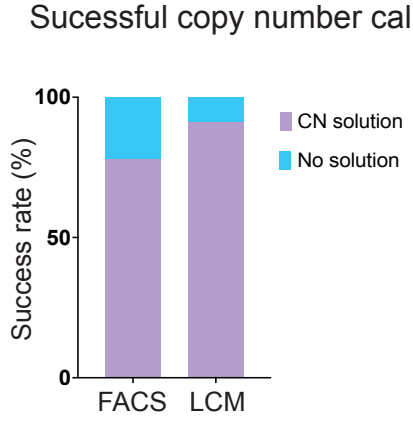


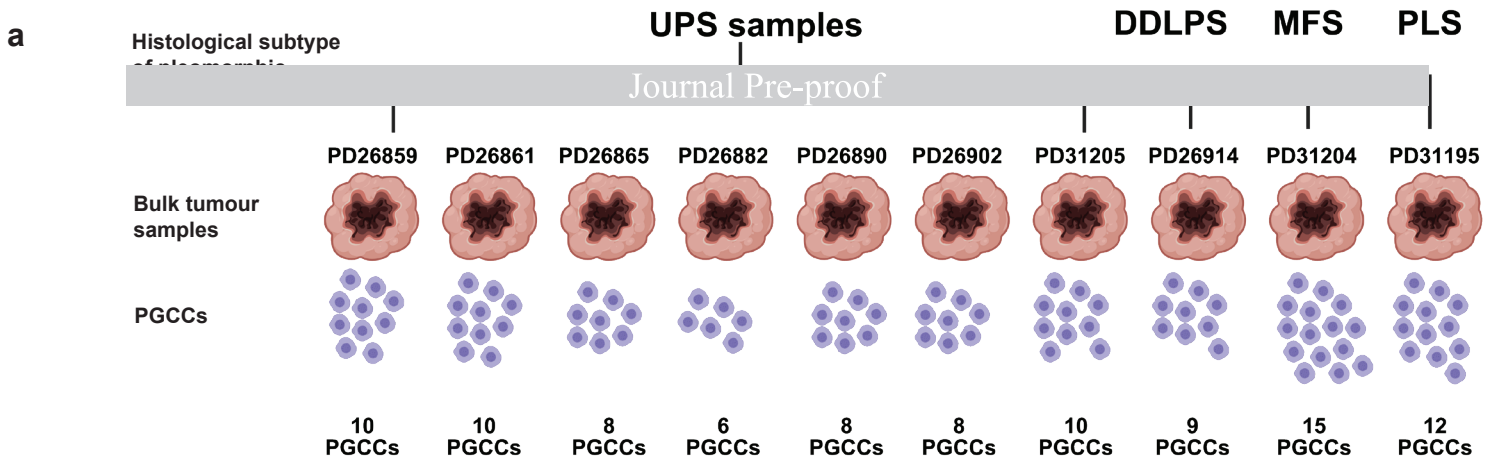


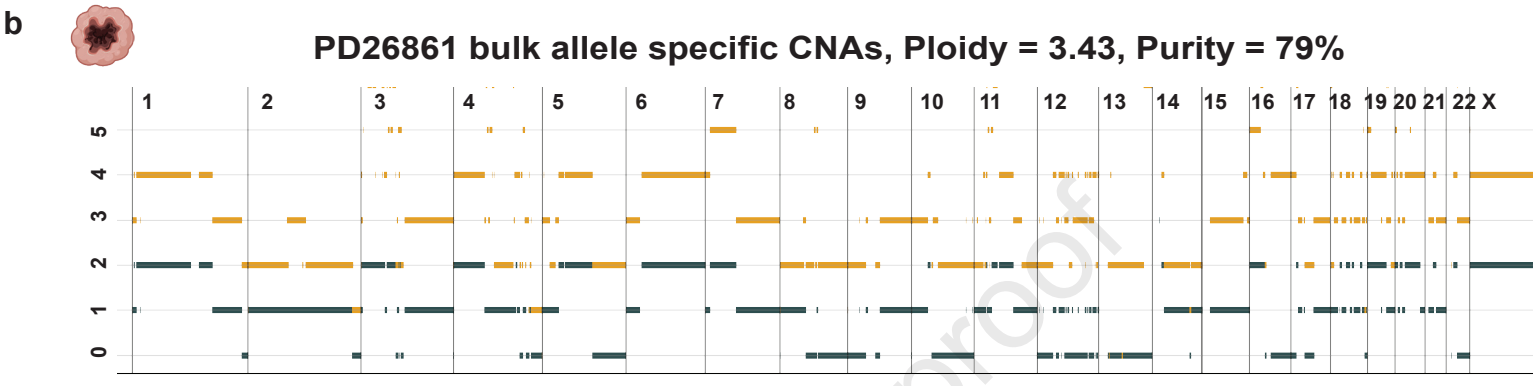


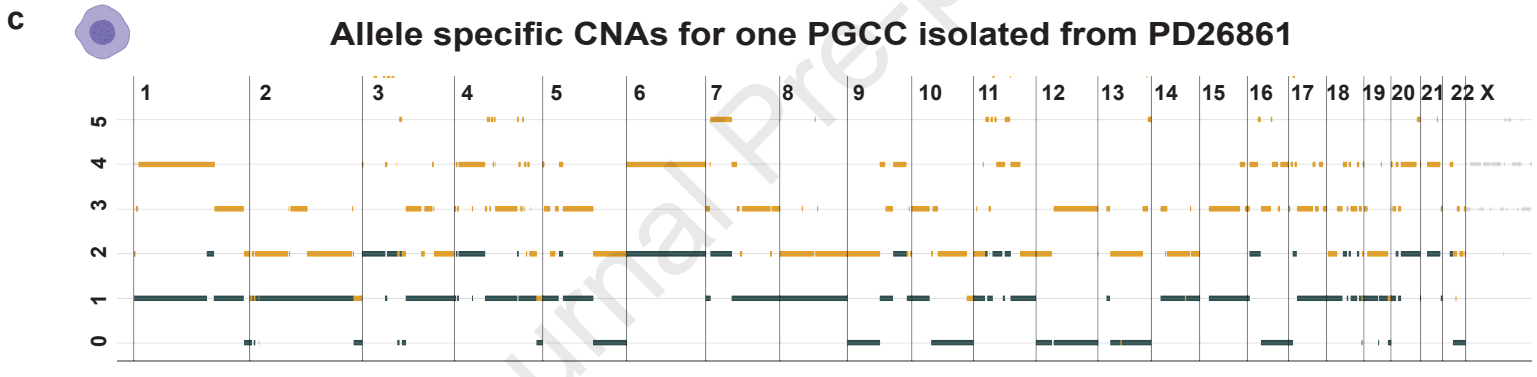


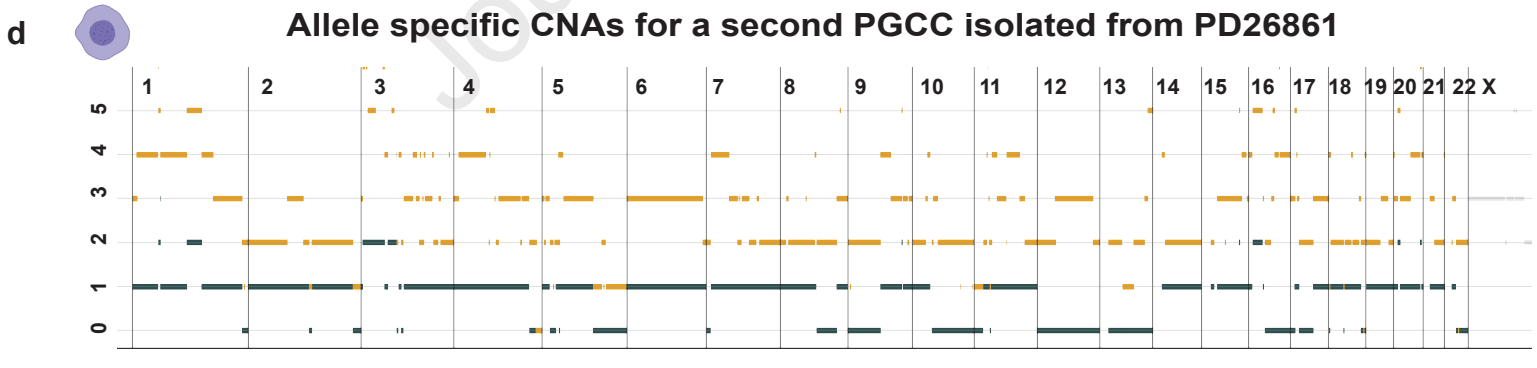


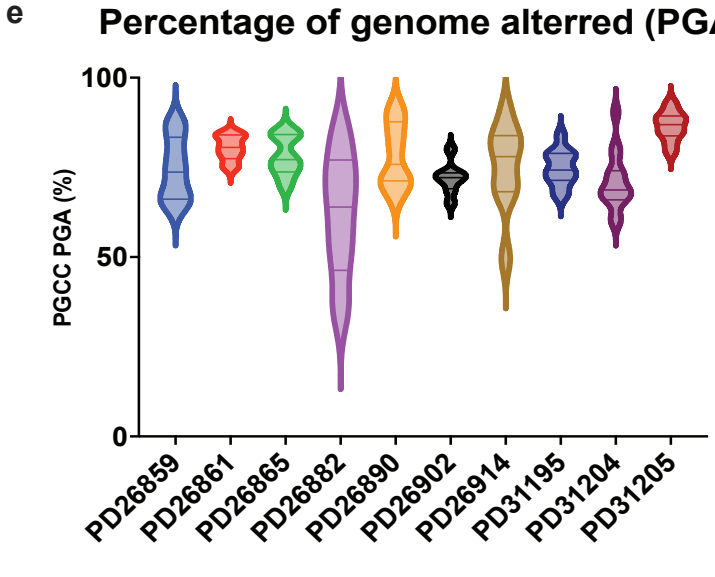


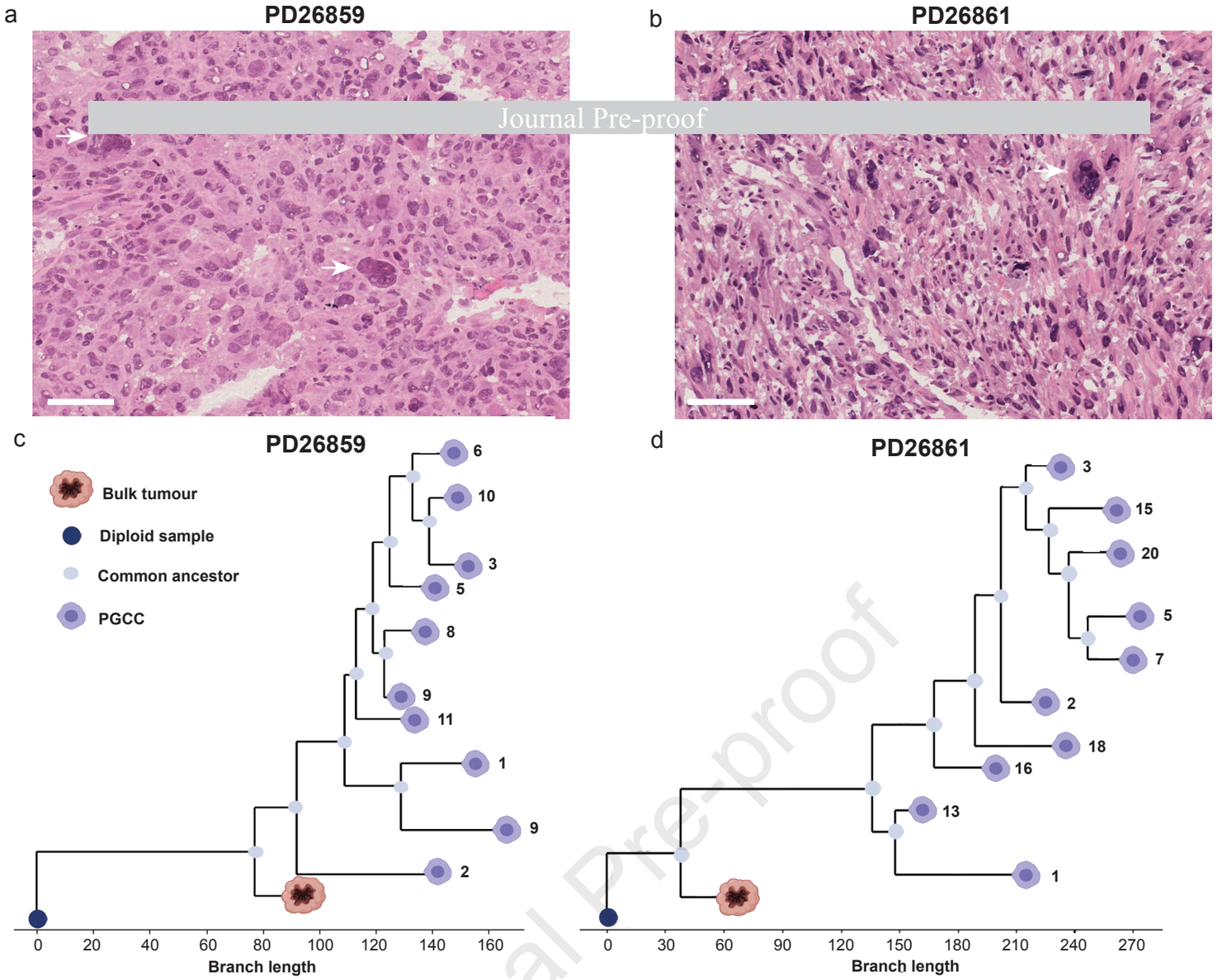


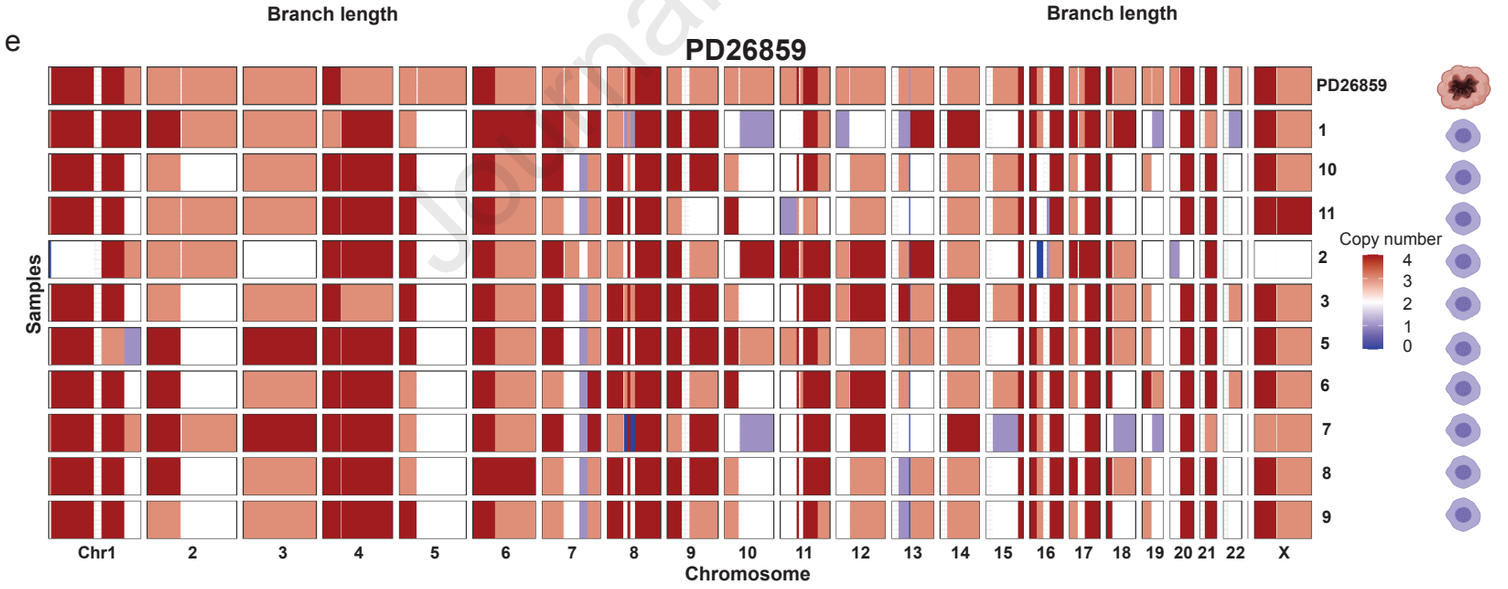


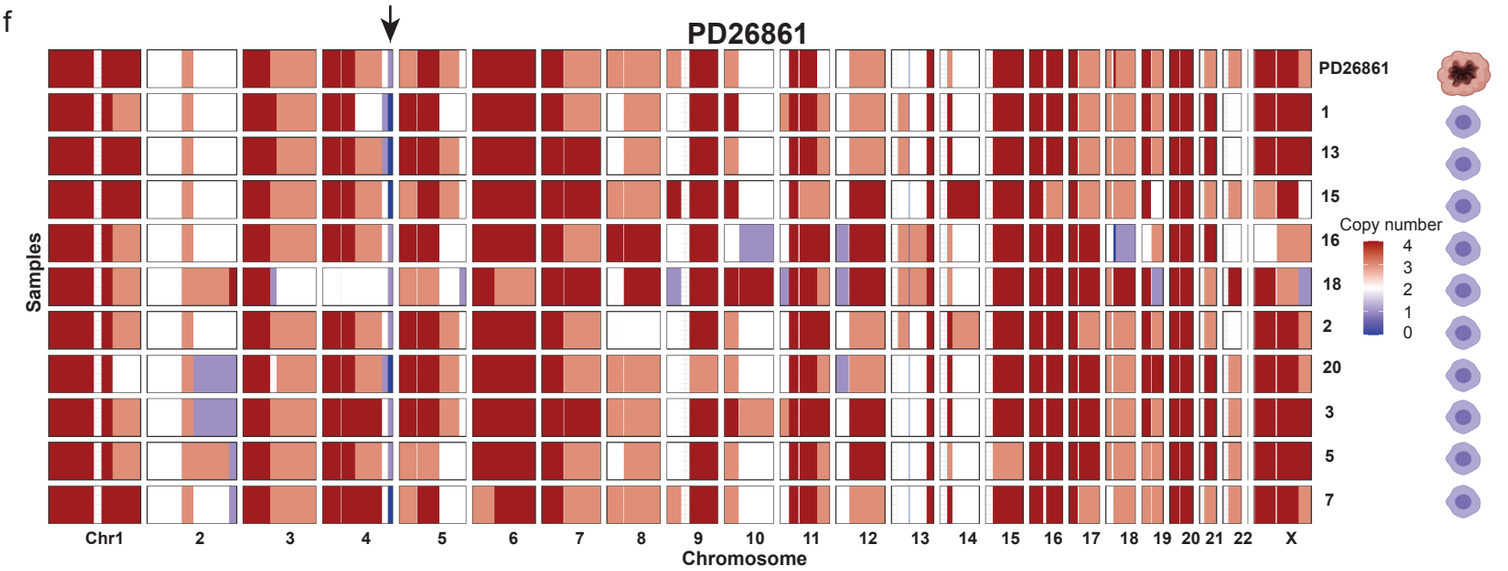


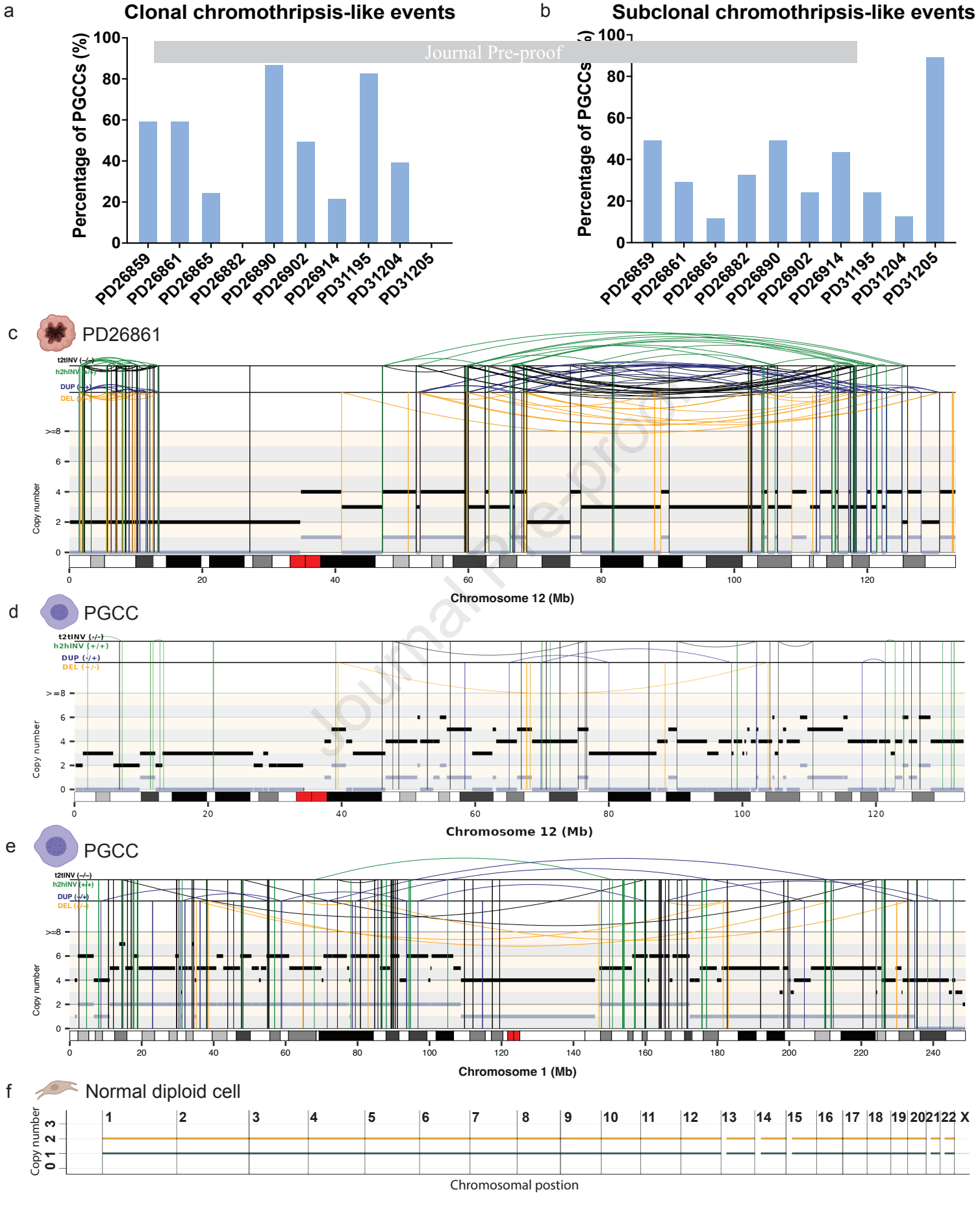




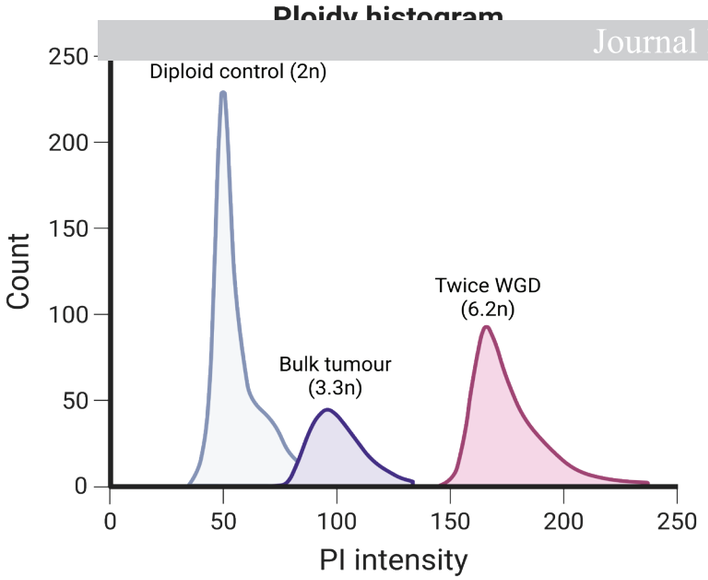




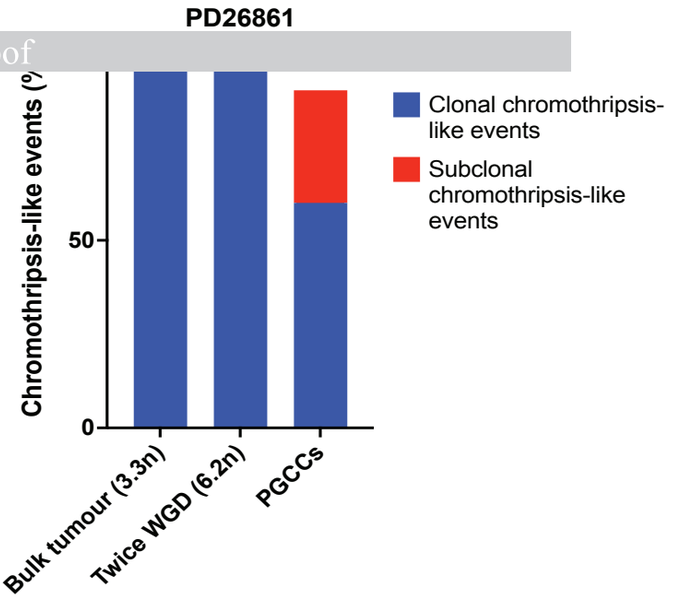




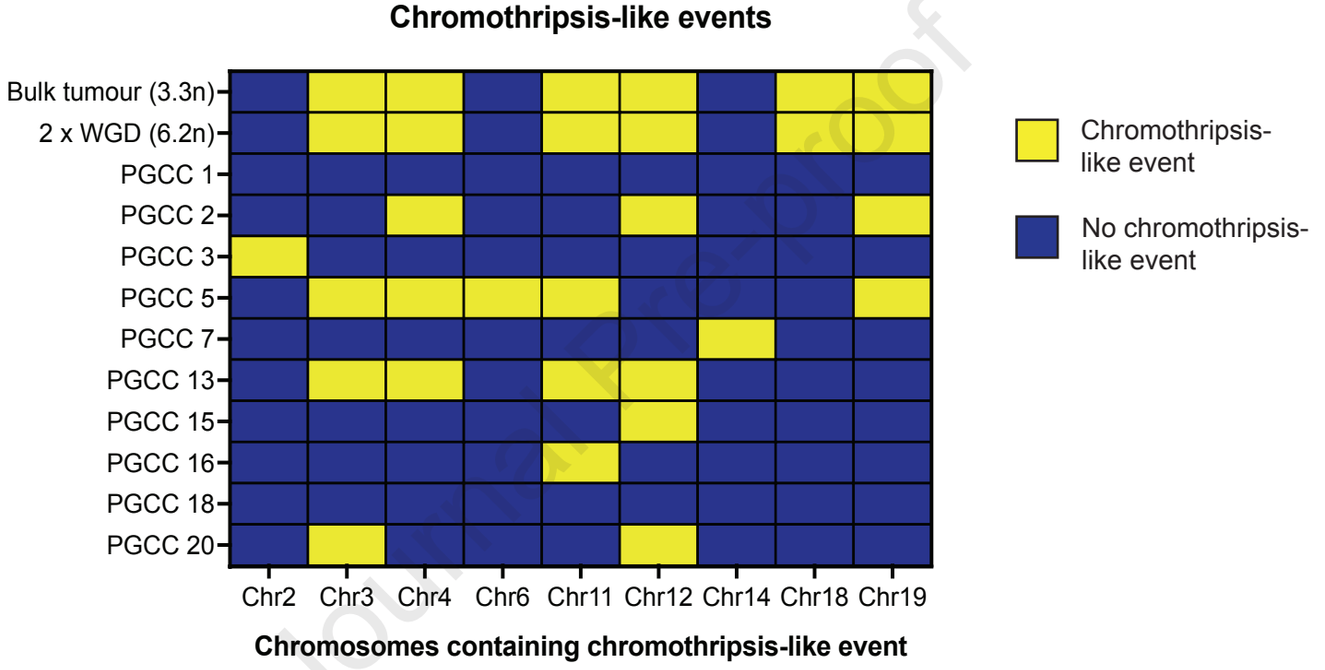
a



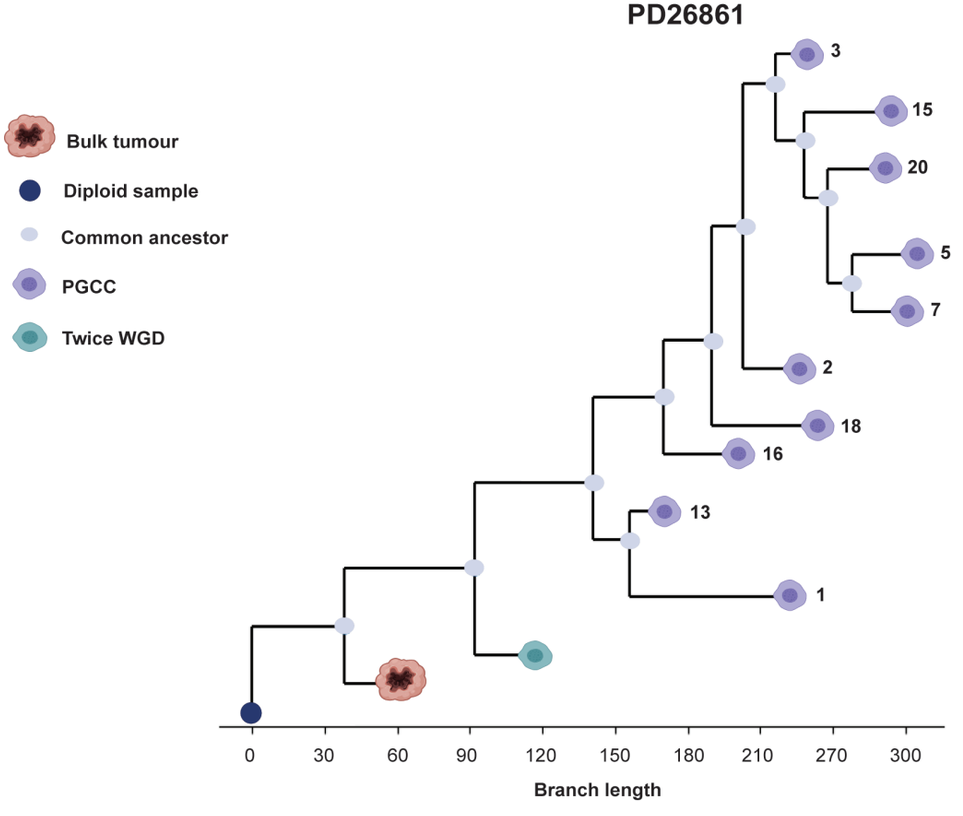
b



c



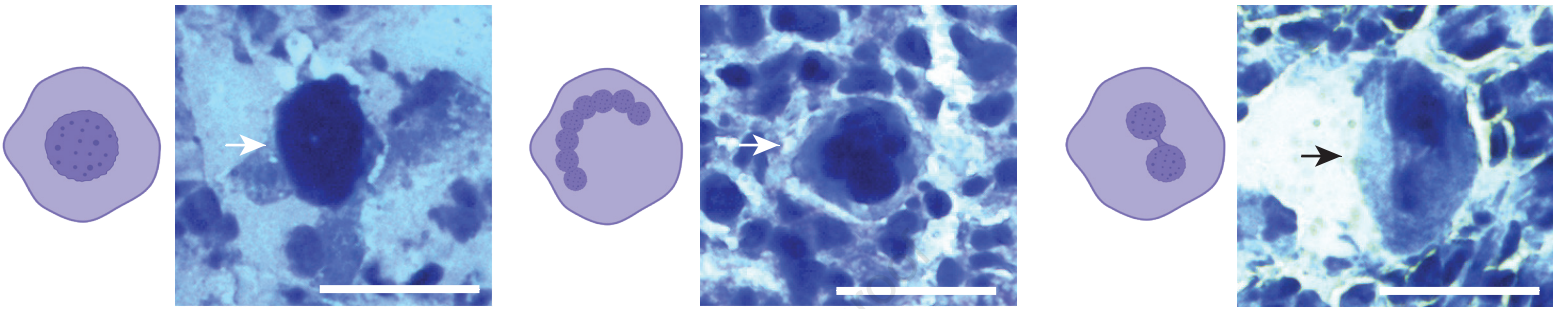
d



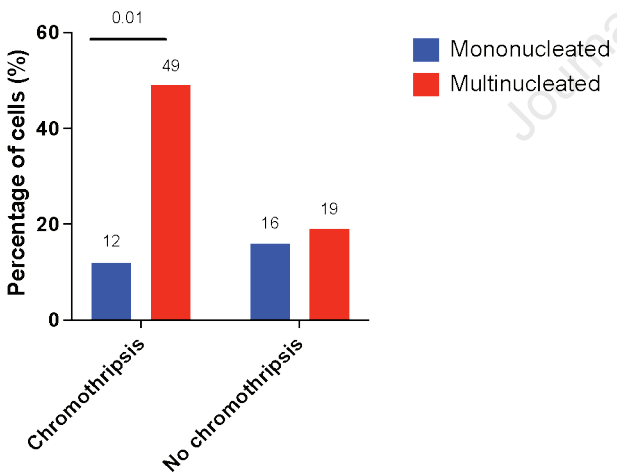
a Mononucleated PGCC

b Multinucleated PGCC

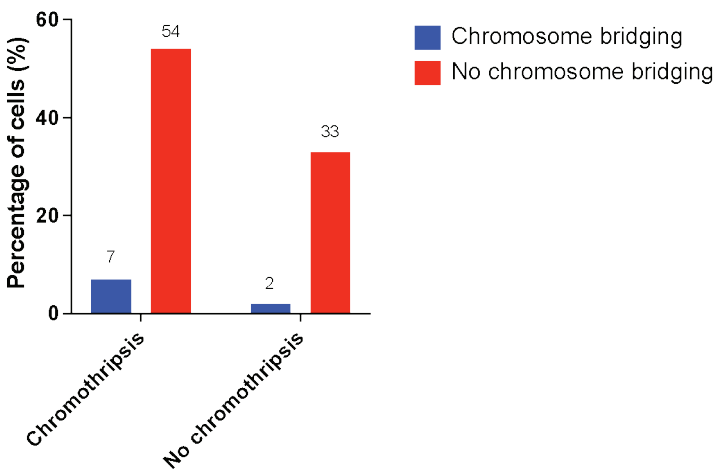
c Chromosomal bridging



d PGCC morphology



e Chromosome bridging



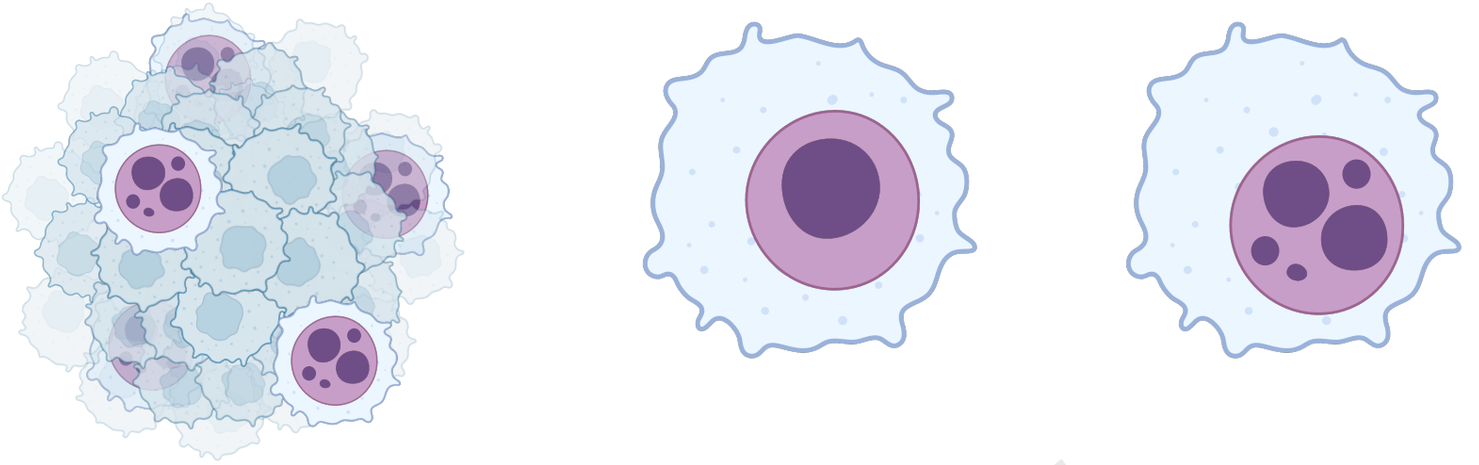
a

Polyploid giant cancer cells (PGCCs)

Tumour tissue

Mononucleated PGCC

Multinucleated PGCC



b

Chromothripsis

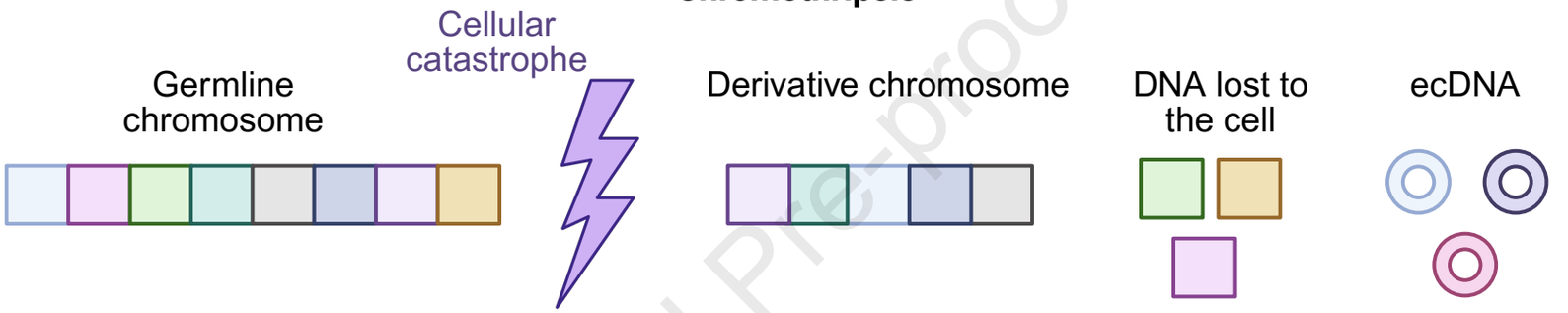
Germline chromosome

Cellular catastrophe

Derivative chromosome

DNA lost to the cell

ecDNA



c

Clonal chromothripsis event

Subclonal chromothripsis event

PGCC

WGD intermediate cell

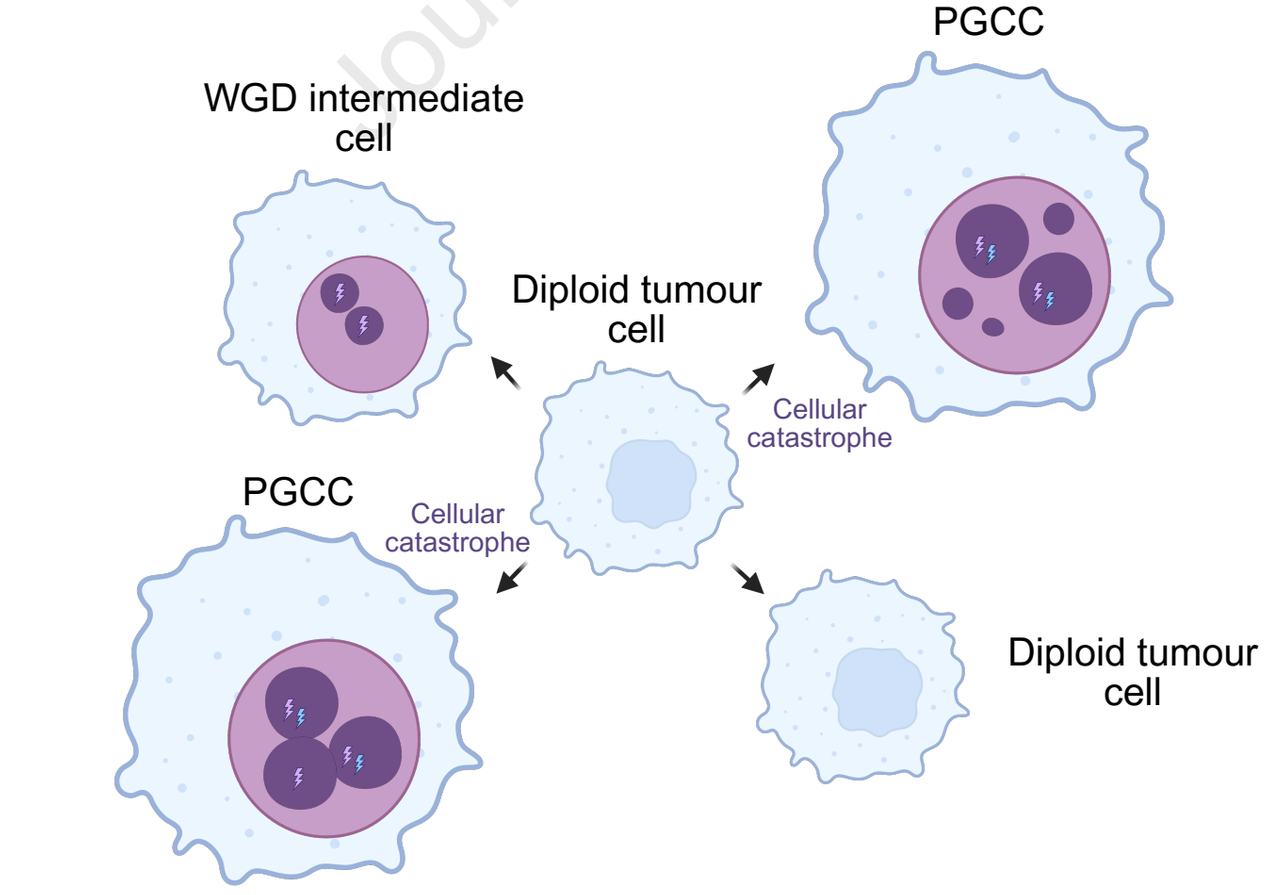
Diploid tumour cell

Cellular catastrophe

PGCC

Cellular catastrophe

Diploid tumour cell



Highlights

Topographic scDNA-seq of 10 UPS samples reveals the genomic traits of PGCCs.

PGCCs show random tumour distribution, suggesting *de novo* formation.

PGCCs arise from the dominant tumour population and display rich CNA heterogeneity.

PGCCs contain clonal and subclonal chromothripsis-like genomic events.

Chromothripsis-like events in PGCCs associate with multinucleation.

Declaration of Interest Statement

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐ The author is an Editorial Board Member/Editor-in-Chief/Associate Editor/Guest Editor for this journal and was not involved in the editorial review or the decision to publish this article.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: